

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: April 10, 2004, 16:23:03 ; Search time 22322 Seconds  
(without alignments)  
17617.209 Million cell updates/sec

Title: US-09-831-000-1\_COPY\_60966\_70038  
Perfect score: 9073  
Sequence: 1 atggagagtcctcgatg.....ggggcaaggatcgcttga 9073

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score  | Query Match | Length | DB ID  | Description          |                      |
|------------|--------|-------------|--------|--------|----------------------|----------------------|
| 1          | 9073   | 100.0       | 133719 | 14     | AF083501 Macaca mu   |                      |
| 2          | 8526.6 | 94.0        | 130733 | 14     | AF210726 Macaca mu   |                      |
| 3          | 1714.4 | 18.9        | 1945   | 14     | AF241163 Cercopith   |                      |
| 4          | 1607.6 | 17.7        | 137508 | 14     | KSU75698             |                      |
| 5          | 1602.8 | 17.7        | 133661 | 14     | U93872 Kaposi's sa   |                      |
| 6          | 1002   | 11.0        | 35100  | 6      | AR065850 Sequence    |                      |
| 7          | 1002   | 11.0        | 35100  | 6      | AR127848 Sequence    |                      |
| 8          | 1002   | 11.0        | 35100  | 6      | AR194750 Sequence    |                      |
| 9          | 1002   | 11.0        | 35100  | 6      | BD190715 Unique as   |                      |
| 10         | 974.6  | 10.7        | 184427 | 14     | EHVU20824            |                      |
| 11         | 847.4  | 9.3         | 108873 | 14     | AF318573 Bovine he   |                      |
| 12         | 812.8  | 9.0         | 108409 | 14     | AF083424 Ateiline h  |                      |
| 13         | 770.4  | 8.5         | 112930 | 14     | XG4346 Saimirine     |                      |
| 14         | 764    | 8.4         | 113027 | 14     | AJ410493 Saimirine   |                      |
| 15         | 714    | 7.9         | 35100  | 6      | AR065851 Sequence    |                      |
| 16         | 714    | 7.9         | 35100  | 6      | AR127849 Sequence    |                      |
| 17         | 714    | 7.9         | 35100  | 6      | AR194751 Sequence    |                      |
| 18         | 714    | 7.9         | 35100  | 6      | BD190716 Unique as   |                      |
| 19         | 708.4  | 7.8         | 7023   | 14     | KSU71367 Kaposi's sa |                      |
| 20         | 679.2  | 7.5         | 60326  | 14     | AY170316 Porcine 1   |                      |
| 21         | 670.6  | 7.4         | 130608 | 14     | AF005370 Alcellaphi  |                      |
| 22         | 667.6  | 7.4         | 73200  | 14     | AF478169 Porcine 1   |                      |
| 23         | 654.4  | 7.2         | 59673  | 14     | AY170317 Porcine 1   |                      |
| c          | 24     | 650         | 7.2    | 2926   | 14                   | AF402655 Human herp  |
|            | 25     | 623.2       | 6.9    | 20766  | 14                   | Y03060 Humine herp   |
|            | 26     | 623.2       | 6.9    | 119450 | 14                   | U97553 Murine herp   |
| c          | 27     | 623.2       | 6.9    | 119550 | 14                   | AF105037 Murid herp  |
|            | 28     | 569.4       | 6.3    | 171096 | 14                   | AY037858 Cercopith   |
| c          | 29     | 565         | 6.2    | 171823 | 14                   | AJ507799 Human her   |
| c          | 30     | 565         | 6.2    | 172281 | 14                   | V01555 Epstein-Bar   |
| c          | 31     | 565         | 6.2    | 184113 | 14                   | M0517 Epstein-Bar    |
|            | 32     | 458.2       | 5.1    | 149696 | 14                   | AF319782 Callitric   |
| c          | 33     | 346.6       | 3.8    | 3445   | 14                   | AF091348 Kaposi's s  |
|            | 34     | 346.6       | 3.8    | 3526   | 14                   | AF091349 Kaposi's s  |
| c          | 35     | 346.6       | 3.8    | 3673   | 14                   | AF091350 Kaposi's s  |
|            | 36     | 217         | 2.4    | 3222   | 14                   | HPU23857 Herpesvirus |
| c          | 37     | 213.6       | 2.4    | 3140   | 14                   | L33971 Cortontail    |
|            | 38     | 200.8       | 2.2    | 5829   | 14                   | D10470 Herpes simp   |
| c          | 39     | 200.8       | 2.2    | 154746 | 14                   | Z86099 Herpes simp   |
|            | 40     | 186.2       | 2.1    | 1772   | 14                   | AY131198 Human her   |
| c          | 41     | 184.2       | 2.0    | 30000  | 14                   | Z48053 Bovine herp   |
|            | 42     | 184.2       | 2.0    | 135301 | 14                   | AJ004801 Bovine he   |
| c          | 43     | 184.2       | 2.0    | 138390 | 14                   | AY261359 Bovine he   |
|            | 44     | 183.8       | 2.0    | 110637 | 14                   | AB024414 Gallid he   |
| 45         | 183.8  | 2.0         | 164270 | 14     | AB049735 Gallid he   |                      |

ALIGNMENTS

RESULT 1

AF083501 AF083501 133719 bp DNA linear VRL 26-JUN-2000  
LOCUS Macaca mulatta rhadinovirus 17577, complete genome.  
DEFINITION  
ACCESSION AF083501  
VERSION AF083501.3 GI:8714565  
KEYWORDS  
SOURCE Macaca mulatta rhadinovirus 17577  
ORGANISM Macaca mulatta rhadinovirus 17577  
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
Gammaherpesvirinae; Rhadinovirus.  
REFERENCE  
1 (bases 1 to 133719)  
Searles,R.P., Bergquam,E.P., Axthelm,M.K. and Wong,S.W.  
AUTHORS Sequence and genomic analysis of a Rhesus macaque rhadinovirus with  
TITLE similarity to Kaposi's sarcoma-associated herpesvirus/human

Pred. No. is the number of results predicted by chance to have a

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

herpesvirus 8  
J. Virol. 73 (4), 3040-3053 (1999)  
99174001  
10074154  
2 (bases 1 to 133719)  
Searles,R.P., Bergquam,E.P., Axtheim,M.K. and Wong,S.W.  
Direct Submission  
Submitted (11-AUG-1998) Division of Pathobiology and Immunology,  
Oregon Health Sciences University/ Oregon Regional Primate Research  
Center, 505 NW 185th Avenue, Beaverton, OR 97006, USA  
3 (bases 1 to 133719)  
Searles,R.P., Bergquam,E.P., Axtheim,M.K. and Wong,S.W.  
Direct Submission  
Submitted (16-JUL-1999) Division of Pathobiology and Immunology,  
Oregon Health Sciences University/ Oregon Regional Primate Research  
Center, 505 NW 185th Avenue, Beaverton, OR 97006, USA  
Sequence update by submitter  
4 (bases 1 to 133719)  
Searles,R.P., Bergquam,E.P., Axtheim,M.K. and Wong,S.W.  
Direct Submission  
Submitted (26-JUN-2000) Division of Pathobiology and Immunology,  
Oregon Health Sciences University/ Oregon Regional Primate Research  
Center, 505 NW 185th Avenue, Beaverton, OR 97006, USA  
Sequence update by submitter  
On Jun 26, 2000 this sequence version replaced gi:5508940.  
Location/Qualifiers  
1. 133719  
/organism="Macaca mulatta rhadinovirus 17577"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:83534"  
repeat\_region  
1. 7840  
/note="partial terminal repeat"  
1353. 2624  
/codon\_start=1  
/product="r1"  
/protein\_id="AAD21330.1"  
/db\_xref="GI:4494908"  
/translation="MFVLVFLMLQPVSVELLPAKLTSTVPCPPHPGDTYLLTCGTT  
STARQDQWFRNTLMRSGNFYGLVSVTPNATISDRYACQTKTITKSNIDFRVS  
SSRLTQERCSGYTYANNTRVLCYSGGNVTRNVLNHWFLANGTAVINGTNNITFPV  
LTKGTGYTCSAFIGNKYSQITINVFFTSFTEKTDIPNEHFKGQIQQTASV  
QHPNYYVFPVPSIGLFGIAISLIMCWLFTRCNSENSESTNYSQTSYIQPSH  
NQRNTNCECRHRYRNAHQESIEELPNOHTSETDSCCOLVLEKVNVAIDPOENTII  
NEWMEQVDDVYVNIHETSIEDVEMHVDYSDTNPENYSGILIEVDVDFNLELN  
QYHGLLDNHNENLHNMIEQIDMLE"  
complement (2692. 3258)  
/note="ORF 2; similar to Kaposi's sarcoma-associated  
herpesvirus ORF 2"  
/codon\_start=1  
/product="dihydrofolate reductase"  
/protein\_id="AAD21331.1"  
/db\_xref="GI:4494909"  
/translation="MDITVNCIVAVDEQLGIGKNGTMPYLYLNEMMYFQKMTSTPSV  
VGEKVVIMGKRTWTFPEKRPVLNRIINILSRELRBPBPHGAFLARTLDDAFNFR  
QYKLKEQLNTVWITGGKSVYESVLNPKLYITRIMESFDCDFVFPSPINFEYTMFL  
SEIFGKDTNPEENGIKYKQVYEKNFK"  
3676. 5613  
/note="ORF 4; similar to Kaposi's sarcoma-associated  
herpesvirus ORF 4"  
/codon\_start=1  
/product="complement binding protein"  
/protein\_id="AAD21332.1"  
/db\_xref="GI:4494910"  
/translation="MTFKLFLHLMYVHGDENCKPPHFTFYRKGSKNDKLSV  
GETPRLCRGYNTKILTECLQNGTWSTPNPCDCKECPADLLGAVHGHGD  
NALKFGNISTECNEGDLGQSNVFCILODTENVNWDSENFCEIQKIKPAVZHG  
DYLPNQVYNGDAITFKCSLYTLVGSITLVCTSNKKWNSPPTCLMLVCESFQIDN  
GYIDIGLSRNYHQQSITVKCSQDYNIVGPETLCTNTTWWPLPKVLVNNPSPFM  
PETPMPETPTDYQKINLSAKTATTENAFVTVSPKDDVT CVKPHFERFVKAEN  
DKKYSVGASVELLCRPGTKMQGSTVEGLSNGTWAPNAKCHRRKCPTELLNGE  
VYVTSGDFAFKYGNITYKNGEYQGLGSMVRCMLDKLTKVDWBPKPAPICEKCK  
PPPOITNGKHPKDYFQYLLDTTFPSCNRDPSLVGDMWTCISNTWNKPPRCEQITC  
SAPNIAHGKLTGSSSVYKYGQSVTICGCTGFTLIGSEISTCKDSSNDPPLFCVPAV

CDS  
6045. 9443  
/note="ORF 6; similar to Kaposi's sarcoma-associated  
herpesvirus ORF 6"  
/codon\_start=1  
/product="ssDNA binding protein"  
/protein\_id="AAD21333.1"  
/db\_xref="GI:4494911"  
/translation="WASKGAPLEDNQSRAPIGACGVYVYAKQDFPFABASILG  
NRPSGVSFSLPILYGLTVEHEPFLTYKAAKVKDVTTLAVKTCFREVTFNNSL  
FRPVDGTGNNELCEARALFGYTFIPGPPHSHWNPDLPCQPLDXDMPGLGVVTE  
GFKERLWRGCLVPAVFTQVQIAGRAFQKVPFLYDEDLFAPGHGRMPFRFVHKQVAYL  
YDSLTSTIAGALRKDVTVIHAKEQFMQDHYKIAIKVAKQFSTPLTKTDDGSSHM  
IVDSVABEALSYGCMFLEQPDACELINDSWPIFDGCDSPPEARVALERWAEQAV  
HVAGOLFAANSVLTKVQKQPRGQKDVNVNSFELQGLGFLNEATIKENGSEAF  
KGVSNALDGSSFTPHLAVAASPSHLAKLCYMQFLOHKKSTNQAFNMVHVGT  
AANSEMTLCHGNTPACLTATLYRKLKDRFPATVTPQRDRPYVVTGTAGTNDLEILG  
NFPSRDREEDNPADDEHPKITYWLCQVTEKLSAIGIEDHNNHNLNINIOSPLR  
VFGIDISIVGEVMKFNMIKNFNFRHVKSVHHILQFCNNYVQWAPCAVFNLLIY  
KSLLIWIIQDICLPYMIYEQNPANGILPSEWVKHGFQTLWTFNFKACLDRGLVTGCE  
LKIIVHRMFCDFDTDAGSGLMAFPFMQVRIARAMVVPKSIKIKRIIFPSNTAGE  
AVQSGFVKPTGTRDYVAGPYMKFLNSLRALPDDTKTAALYLWHKISQTNKTPVLK  
DVDPDELAELSVYKTNLSLAPEETNVDVDPDSLSVARIKNGAILRACQIQFYAT  
TLHCLTVPQLIDAEYPHVLGSAALATPVAYLAEIRGTALTQVTTARQPVAAATGR  
RPLVITVMVNNKTYGVNGNNVNFHCGNIGHPAGRGVDRNMLWPSSPFKKTGVSAMLK  
RHWMTFIIDELIKRAAQGTISTFEAESVKRSVQALLDKDNPLLSKSVILELRHLG  
KQCQDLSSEDDVYGLDYCMLTDEVLFTLDNIAQSGVPTWIEDAGALIEDHQDADDLQ  
FVDSDDIATACQPEEQPLTPPSAGALLAGKRXKINALLSLDLI"  
9468. 11528  
/note="ORF 7; similar to Kaposi's sarcoma-associated  
herpesvirus ORF 7"  
/codon\_start=1  
/product="transport protein"  
/protein\_id="AAD21334.1"  
/db\_xref="GI:4494912"  
/translation="WARELAALYAQLSALAVDSLIVFADPRSIDGARILKTKTQIEN  
LNEDLLPLLRQNSVEVSSLSLEVEHLAKNEIDKGLERLQRYSSREHFEVLHLR  
PECHYHSTVTTPGGGLIDVNMCLINDVELLCKELGSVFYICIGANALSGLENLVLF  
FECYHSTVTPHPPDLVTVSPCVQCLRIELIPLVQGSLLAVLADHCDHLCKVRAE  
PIHGLFELSGLQKVTKRSDDATQGVRRSSDQLRSSLAALQDNHIFKRVASIME  
LSNTIYNNAGTGLQGTENECQMARLLTHEADHERRALITPLKATHYDFCRPD  
PISELCGLFNSIDDTINALRSCSVTFQQAANYNMRKQNELFELNSILRQSGSA  
GSOQPATSPERTTVAATAASDVIKQAEQYMKVARDGFKLCTEQIQAVAL  
ANALCMRVGCVAYGEASELVNHLRERFVALPWEARCRSDQILFENSKIKNSLYS  
QRLREHVEIITLQFYGLITGLRQSDLPFGPANVALAOCFEAGMLPHKMLVSEM  
IWPQIQPKWDIDQTFNRFYQIPGGLNNAQKAMCFIRELVSLVALNRTWKTIRIF  
SLAREKLSISNLDVKLTSGLYLTYEQDAPLVLLISQNTGWIIFKDLIALLYLHLQUSDG  
HDDN"  
11515. 14004  
/note="ORF 8; similar to Kaposi's sarcoma-associated  
herpesvirus ORF 8"  
/codon\_start=1  
/product="glycoprotein B"  
/protein\_id="AAD21335.1"  
/db\_xref="GI:4494913"  
/translation="MMITNTRRLRLRAWVIIAIGTAVGENVTTPKGATTAKTPPGP  
STTPPENPRAEAFKPRVCASATGELFRFNLEKTCPTGEDTKHGEGLIMVFKNIV  
PHIKVRYRVATSVTVYRGWTSTAVTGKQEVIRPVQYIEINMDTITYQCFSSNRVN  
VNGIWTYTDSDFTNQVFLQVBSGLTDNIQRYSPQVLYTTPGWPGIYRVRTVNC  
EIVDMIARSAEPYIFVYFALGDTVEVSPFCNDSTCSVAETENGLGARVLTNTIVD  
FATQPTTETRVFADSGSYTVSKAEDPKSAVCALTLMKTFPRAIQTHESAYHFVAN  
DVATFTSPLSQVNTFTGYTFCINDIQTLNATIKLSDTHATNGSQSYETEGGLF  
LMWQPLSLADEMRRLNGTTPAPTITSTANRVRSVGTNEQATDLDAAAPQOFAY  
DKLRSINKLLEELSRACREQRDVTMYELSKINPTSVMTAYIGREVSAXFVGDAI  
SVTDCVAVQASVTHKSLRSTPGICYSRPPVTFPLNFTPLFKQLGSPRNEIILTD  
NOVACKTCHSYFIASNVYTYKDYVFKKINTSEISTLGTFIALNLSFTENIDFRV  
TELVSRAKSLSGSVDFDIETNFRYNYTQBLAGLRDLDLTLNRLRLADLSEIV  
ALDVGVTGVNVASSVITLFGSIVSGDGFINTIKPFGGKMLIIVAVLVLEALNR  
TNAQASDLIRMIYDIDKMQSGGKVDGQFQEQEYEMSK"  
14122. 17166  
/note="ORF 9; similar to Kaposi's sarcoma-associated  
herpesvirus ORF 9"

```

herpesvirus ORF 9"
/codon_start=1
/product="DNA polymerase"
/protein_id="AAd21336.1"
/db_xref="GI:4494914"
/translation="MDFPNYLGPRGPRHSHRGTADAPAGAGAVQPPPPVCRLLIPA
KAYDPAGMIFPTIPFPPTFYENGARGDVLNANERSMWTADRPAPVAPDPQDSITF
HYDVVETTHIAADRCAPSPRFQDIIIPSGTIVKLLGRTDGTSCVNVFRQVIFYA
KVPAGINVTHIIQQALKNAGAACGFSRRYKRIKLTIDVAEHPVTEITLSSGSMF
STLSDLVACGGEVESNDVAFRRFVLDHGTFTFGWYSCARATPRLAARDARTALEFD
CNWEDLSVQADSDWPPYRIVAPDIECTGEAGFPCCATRDGDAVIOISCVFVYTRGAP
NPNILFSVGTCDPIIDTDLBFPSEYDMLVSFAMIRDFEDVDFLTGYNISNFDLPYL
IIRASOVNLRNLNETHKIGTSIFVHEPRGGGGGFMRSVSKIKIAGIPIVDIMQVCR
EKLSIDYKLDVAROGLCKKEDVSIDIPFPRSGPGRKAVGSGVCMDSVLVNDL
KUMFTIHDEISIAKIAKIOARVLTDGQQLRVFSLLEAAARENPILPVPPEGGG
YOGATVINIPGFYDEPVLVDVFPALPSYIIQAHNLCTVTHGSDLHLNPLNPDYD
ETPVLGGVHFVFKHRESLRLITVWLEKRRAIRRTLAACDDPFSKLTLDKQOLA
KVTICNAVGTGFGAGLPCINIAETVTLRGKTMLEMSKVVEALTTEDLRLRGRE
IVARHGRFVVVGTDSLFIACDGVSAEAVAFCDLAAIRITADLPFPIKLBKKT
FKCLLLTKKRIYGLVNDKMYMKVGLIRKACKFVQERCBAILDLVHLPEVKARA
RLCKRPHAVTEELGEPAGFIKIVELVINASYLDLNSVVPFIEQLTFSTELSRPVCYK
TNLPHLAVYKLAESCELPQVHDRIPIYFVDAFSGSLKSDLAERHDFVYRHOQIPVAV
DYDFDKLVHGAANILQCLFGNNADTTVAILYNFLNVPYKLS"
17261..18511
CDS
/notes="ORF 10; similar to Kaposi's sarcoma-associated
herpesvirus ORF 10"
/codon_start=1
/product="unknown"
/protein_id="AAd21337.1"
/db_xref="GI:4494915"
/translation="MLVNELSVLGDWVTFHRRFSFVNLRLQTFKRGHGYARVRL
PRLDOLHGHAFGLVTRKELPPSPDCVALIAPLDGGSDADAAVAPGFVLDSSRP
LVTWNNAGHTIRFCLLPKLDLERAVTVFVNGSGASEGPKPTKATESLPGGP
LAVSGASQTSFHSFVAYPPTANSVACLILRLQVRPSPDDAAHDAHRLSKPYVFSN
SGNVCKASHVTLSPBRCTAQMEIIFAGDQNAEIVLQSGFVLPHTHGRVLYA
DAEKTQPGSSAAYVOLIFQQAARGDLAFVTVGAPEPLFVTVPALLSGGCTHLL
RLFPNGTPTTKRDLTVAAAPCPVVRLLSADDAADPRDLVASPDTGALSINAFITPVG
FPGVSAECHVSLRDNGVHERMNH"
18520..19749
CDS
/notes="ORF 11; similar to Kaposi's sarcoma-associated
herpesvirus ORF 11"
Query Match 100.0%; Score 9073; DB 14; Length 133719;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGAGTTCCTCGGATGGACCAACACGTCGACCAACATCCGGGGTTCATCTTGAC 60
DB 60966 ATGAGAGTTCCTCGGATGGACCAACACGTCGACCAACATCCGGGGTTCATCTTGAC 61025
QY 61 ATGACGTCGATGCCAAAGTCAGGGGTGTGCTGGATCAAGTCAGTCCGCTGTCAAATATA 120
DB 61026 ATGACGTCGATGCCAAAGTCAGGGGTGTGCTGGATCAAGTCAGTCCGCTGTCAAATATA 61085
QY 121 ACTACAGCCACCGGAATGGGTGTGATGACGACCTGGCCTTCGATCCGGCTGAAGACTCC 180
DB 61086 ACTACAGCCACCGGAATGGGTGTGATGACGACCTGGCCTTCGATCCGGCTGAAGACTCC 61145
QY 181 GGGCCGTTCTGCGTTTACGGTTTATCTAATACGGGAACCTGCTGGTGTGGAAAGT 240
DB 61146 GGGCCGTTCTGCGTTTACGGTTTATCTAATACGGGAACCTGCTGGTGTGGAAAGT 61205
QY 241 ACCAGCATATCGGCCCTGTACCAAAATTTAAATCTGCTGATACAGGGGCGGACCAACATA 300
DB 61206 ACCAGCATATCGGCCCTGTACCAAAATTTAAATCTGCTGATACAGGGGCGGACCAACATA 61265
QY 301 GCCGCACAGAACTATCGCTCGCCTAAAGACGTTCTGTCCACGATCTTCAGCGGTTTT 360
DB 61266 GCCGCACAGAACTATCGCTCGCCTAAAGACGTTCTGTCCACGATCTTCAGCGGTTTT 61325
QY 361 GGGTTTTAAGAGCGGACACATATATAGCGGTGAGAAAAGCTCATCAGACCGGAGCCGTA 420
DB 61326 GGGTTTTAAGAGCGGACACATATATAGCGGTGAGAAAAGCTCATCAGACCGGAGCCGTA 61385

```

```

QY 421 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTGGCCGGTTATAGTCGACATT 480
DB 61386 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTGGCCGGTTATAGTCGACATT 61445
QY 481 ATGAAAAGAGTATTGGCGAAAAAACCCTTGTATCTACCTACCGCTCGTCTTTTACTGCTTC 660
DB 61446 ATGAAAAGAGTATTGGCGAAAAAACCCTTGTATCTACCTACCGCTCGTCTTTTACTGCTTC 61505
QY 541 TTTGAAAACCCCTCTCGAGAAATGACCGGACCGGTGTTTATGGAGCTTCCAATATTATTGTAATC 600
DB 61506 TTTGAAAACCCCTCTCGAGAAATGACCGGACCGGTGTTTATGGAGCTTCCAATATTATTGTAATC 61565
QY 601 GACGAGCGCGAAACCCCTGTCCTCTTATATATCTACCTACCGCTCGTCTTTTACTGCTTC 660
DB 61566 GACGAGCGCGAAACCCCTGTCCTCTTATATATCTACCTACCGCTCGTCTTTTACTGCTTC 61625
QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGCGGCGTTCCTGTCATAGTAGTC 720
DB 61626 CTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGCGGCGTTCCTGTCATAGTAGTC 61685
QY 721 GTCGGTTCGCAACGCGAGCAAAACCGGTTCCAGTCACTTACCAACCGGAGCGGACGCAAAAG 780
DB 61686 GTCGGTTCGCAACGCGAGCAAAACCGGTTCCAGTCACTTACCAACCGGAGCGGACGCAAAAG 61745
QY 781 ACGGAGATATCGTCTGCGAGAACATCTTAACATTCATGATAGGGAAGAGTCTGTATCT 840
DB 61746 ACGGAGATATCGTCTGCGAGAACATCTTAACATTCATGATAGGGAAGAGTCTGTATCT 61805
QY 841 GAGTACGTACACTTTGGAGAGAAACTGGGGCGTGTATTATAAACAACAAGCGGTGACCGAT 900
DB 61806 GAGTACGTACACTTTGGAGAGAAACTGGGGCGTGTATTATAAACAACAAGCGGTGACCGAT 61865
QY 901 CTGCAGTTCGACACCTCTTAAATAATTTTATAGATATATCTTCCCATTTCTTACGAGAGTC 960
DB 61866 CTGCAGTTCGACACCTCTTAAATAATTTTATAGATATATCTTCCCATTTCTTACGAGAGTC 61925
QY 961 ATGAGTTACGTAGACAGATTTGTCGTTCTTAAAGTAAAGTATATGATTCCTTTAGATATAC 1020
DB 61926 ATGAGTTACGTAGACAGATTTGTCGTTCTTAAAGTAAAGTATATGATTCCTTTAGATATAC 61985
QY 1021 ATTGGTGGACCCGACTCTTTTGTTCATAGCGAGGTAAAGCGTATCTTAACAAACCTA 1080
DB 61986 ATTGGTGGACCCGACTCTTTTGTTCATAGCGAGGTAAAGCGTATCTTAACAAACCTA 62045
QY 1081 CACACATGCTTAAACGCTAGGGGGCGATACCAAGGACACGAGCTCTTTTACCTGTCCCGTG 1140
DB 62046 CACACATGCTTAAACGCTAGGGGGCGATACCAAGGACACGAGCTCTTTTACCTGTCCCGTG 62105
QY 1141 GTGTGCGAGGTGTTTGTGAAGCCGTTTGAAGATPACAAAGCGGCGGCTCAACCTCACCAAC 1200
DB 62106 GTGTGCGAGGTGTTTGTGAAGCCGTTTGAAGATPACAAAGCGGCGGCTCAACCTCACCAAC 62165
QY 1201 CTCACCGTGCAGGAGTGGGTGACAAAAAACCCTTTTAAAGTTAAGTAAATTTATTCGAGTTT 1260
DB 62166 CTCACCGTGCAGGAGTGGGTGACAAAAAACCCTTTTAAAGTTAAGTAAATTTATTCGAGTTT 62225
QY 1261 GTGGACCAAGGACATGTCTCATAGTCGCCACCGGAATCCACTGAAGCTTCGACCGAGTTACC 1320
DB 62226 GTGGACCAAGGACATGTCTCATAGTCGCCACCGGAATCCACTGAAGCTTCGACCGAGTTACC 62285
QY 1321 TTTATCCCAAGTTTGTGTAAACAGCCACCGCTCTTTTAAATGAAAAAACAACAAAAATGC 1380
DB 62286 TTTATCCCAAGTTTGTGTAAACAGCCACCGCTCTTTTAAATGAAAAAACAACAAAAATGC 62345
QY 1381 ATATGTTGGTTTCAGGGTACGTTACTTTTGTAGTTCAAAAGAACTCTTAGACAGCGAACTCTTC 1440
DB 62346 ATATGTTGGTTTCAGGGTACGTTACTTTTGTAGTTCAAAAGAACTCTTAGACAGCGAACTCTTC 62405
QY 1441 GTGGAGACGCAATTCGCAAGACCGTCCGGAATACGTTATACGTTTCTTAAACACCGTACTG 1500
DB 62406 GTGGAGACGCAATTCGCAAGACCGTCCGGAATACGTTATACGTTTCTTAAACACCGTACTG 62465

```

|    |       |         |                    |                     |                              |                 |       |
|----|-------|---------|--------------------|---------------------|------------------------------|-----------------|-------|
| Qy | 1501  | TACAA   | CGCCATGTACTCGTTTCA | CGCGTAC             | CGCGTGTACTAGGTCA             | CACGAGAAATATCTG | 1560  |
| Db | 62466 | TACAA   | CGCCATGTACTCGTTTCA | CGCGTAC             | CGCGTGTACTAGGTCA             | CACGAGAAATATCTG | 62525 |
| Qy | 1561  | CAAGAC  | CTAAATTTGCA        | CCCTCCGCGCTCTG      | CAACCGCGCGTGTAGAC            | CTTCAA          | 1620  |
| Db | 62526 | CAAGAC  | CTAAATTTGCA        | CCCTCCGCGCTCTG      | CAACCGCGCGTGTAGAC            | CTTCAA          | 62585 |
| Qy | 1621  | ACGGTT  | CGTGAAGAGTTAAAC    | CTGGAAGACACATCTT    | CTCAACCGTGTGAC               | CTCCCGG         | 1680  |
| Db | 62586 | ACGGTT  | CGTGAAGAGTTAAAC    | CTGGAAGACACATCTT    | CTCAACCGTGTGAC               | CTCCCGG         | 62645 |
| Qy | 1681  | CCAC    | CGCGGGTATCAC       | CTCCCTCCAGGTTTTGT   | CGACACGCTAC                  | CGCCCTTAAGGAC   | 1740  |
| Db | 62646 | CCAC    | CGCGGGTATCAC       | CTCCCTCCAGGTTTTGT   | CGACACGCTAC                  | CGCCCTTAAGGAC   | 62705 |
| Qy | 1741  | GTCT    | CGCTCCAGAAATAAAG   | TGCGCTGTCTGCTGGTTTG | CGCGGAGTTTGAGAAAGGAA         | 1800            |       |
| Db | 62706 | GTCT    | CGCTCCAGAAATAAAG   | TGCGCTGTCTGCTGGTTTG | CGCGGAGTTTGAGAAAGGAA         | 62765           |       |
| Qy | 1801  | ACGTTTT | CCGGTTTACGGTTAAAC  | TGCTGTAAGGACGAGTTG  | ACTTTGTTCTCCCT               | 1860            |       |
| Db | 62766 | ACGTTTT | CCGGTTTACGGTTAAAC  | TGCTGTAAGGACGAGTTG  | ACTTTGTTCTCCCT               | 62825           |       |
| Qy | 1861  | TCAGAA  | CGCTCTCAAC         | CGGGCTGTGCGGTTTG    | CTCGACCGTTGAAATCTATAAAATTAAG | 1920            |       |
| Db | 62826 | TCAGAA  | CGCTCTCAAC         | CGGGCTGTGCGGTTTG    | CTCGACCGTTGAAATCTATAAAATTAAG | 62885           |       |
| Qy | 1921  | GGTA    | CACGTTTTTACCGGTAG  | CGTTCCGTCGTGTCA     | GGGTTTGCCACTCAGTGATGAC       | 1980            |       |
| Db | 62886 | GGTA    | CACGTTTTTACCGGTAG  | CGTTCCGTCGTGTCA     | GGGTTTGCCACTCAGTGATGAC       | 62945           |       |
| Qy | 1981  | CTCAG   | GAAGAAGATGCCCTCC   | CTGCTGTCAGGAGCTCT   | AGCGGTTTTATCGCGTGCTTA        | 2040            |       |
| Db | 62946 | CTCAG   | GAAGAAGATGCCCTCC   | CTGCTGTCAGGAGCTCT   | AGCGGTTTTATCGCGTGCTTA        | 63005           |       |
| Qy | 2041  | GAGA    | ATAACATAA          | CCAAATTCAGAAACCA    | TGAGGAGCGGAGCGTTTTCC         | 2100            |       |
| Db | 63006 | GAGA    | ATAACATAA          | CCAAATTCAGAAACCA    | TGAGGAGCGGAGCGTTTTCC         | 63065           |       |
| Qy | 2101  | TGTC    | CGGGGACATATCGG     | GTGACCTCAAAATTTAG   | CCATGACCATCTGTAAGGACACAGGGA  | 2160            |       |
| Db | 63066 | TGTC    | CGGGGACATATCGG     | GTGACCTCAAAATTTAG   | CCATGACCATCTGTAAGGACACAGGGA  | 63125           |       |
| Qy | 2161  | ATGTC   | GTGGAGCGAGTAC      | CGCTAGTATTTGGGTC    | CCACAAGACGTC                 | 2220            |       |
| Db | 63126 | ATGTC   | GTGGAGCGAGTAC      | CGCTAGTATTTGGGTC    | CCACAAGACGTC                 | 63185           |       |
| Qy | 2221  | GTGT    | ATGTAGCAATATCA     | AGGCGTGTAACTCAAA    | TTATTTGTCATGGACAGAACCC       | 2280            |       |
| Db | 63186 | GTGT    | ATGTAGCAATATCA     | AGGCGTGTAACTCAAA    | TTATTTGTCATGGACAGAACCC       | 63245           |       |
| Qy | 2281  | CTTAA   | AACCTCTCAGAGAA     | CCAGTCGATACACCT     | CCGCAAGCATATAGTCCGCGCC       | 2340            |       |
| Db | 63246 | CTTAA   | AACCTCTCAGAGAA     | CCAGTCGATACACCT     | CCGCAAGCATATAGTCCGCGCC       | 63305           |       |
| Qy | 2341  | CTCA    | CAACCCAAACACAA     | CCCTCATCTAATAATAA   | ATAAACAAGCAATGGA             | 2400            |       |
| Db | 63306 | CTCA    | CAACCCAAACACAA     | CCCTCATCTAATAATAA   | ATAAACAAGCAATGGA             | 63365           |       |
| Qy | 2401  | ATTGT   | GTGTTTTATTTCAGT    | CCAAACCGGACGATGTTG  | CTATTTCCACACCGGG             | 2460            |       |
| Db | 63366 | ATTGT   | GTGTTTTATTTCAGT    | CCAAACCGGACGATGTTG  | CTATTTCCACACCGGG             | 63425           |       |
| Qy | 2461  | GGGT    | CGCGCCATAGTTTGA    | CGACAGAAACCGCCCGG   | CCCTCTGGGTTAGAAGTAGCC        | 2520            |       |
| Db | 63426 | GGGT    | CGCGCCATAGTTTGA    | CGACAGAAACCGCCCGG   | CCCTCTGGGTTAGAAGTAGCC        | 63485           |       |
| Qy | 2521  | CGCC    | CCCGTGTGTAGGTCAT   | AGCTCTCTTAGGSCCT    | CTGGGTTGTAACCAATAATCA        | 2580            |       |
| Db | 63486 | CGCC    | CCCGTGTGTAGGTCAT   | AGCTCTCTTAGGSCCT    | CTGGGTTGTAACCAATAATCA        | 63545           |       |
| Qy | 2581  | CGTC    | GTATCACCGGTTG      | CGAGTCCCTCGTGTG     | ACGCGGCTCTTTGAGGCGGAC        | 2640            |       |



|    |       |   |       |
|----|-------|---|-------|
| Dd | 64626 | GAGCAGGATGCTCTGTCGGTTTATCAGCGTCGCCCTTGCTATAGCCCTTCTTCCCA    | 64685 |
| Qy | 3721  | CAGCATAAAAACGCACTCTGAAGCTTGGCAGAAAGCGAGCTTATATAGCTGGTGA     | 3780  |
| Dd | 64686 | CAGCATAAAAACGCACTCTGAAGCTTGGCAGAAAGCGAGCTTATATAGCTGGTGA     | 64745 |
| Qy | 3781  | CCAACTCCAGCAGGTTGCGGTGACCCCGCTTCCCTCTTCCACCGTCAGAAATGGT     | 3840  |
| Dd | 64746 | CCAACTCCAGCAGGTTGCGGTGACCCCGCTTCCCTCTTCCACCGTCAGAAATGGT     | 64805 |
| Qy | 3841  | GTTTAAAGCAGAACTCCCGTTTTCGCCAACAGTCTAAGCACCCCGTGAGAGGCGGT    | 3900  |
| Dd | 64806 | GTTTAAAGCAGAACTCCCGTTTTCGCCAACAGTCTAAGCACCCCGTGAGAGGCGGT    | 64865 |
| Qy | 3901  | GAACCCAGGTACAGTATGCTATCTTTTAAATAATATTTTGGAGCTTGGAGGTATCT    | 3960  |
| Dd | 64866 | GAACCCAGGTACAGTATGCTATCTTTTAAATAATATTTTGGAGCTTGGAGGTATCT    | 64925 |
| Qy | 3961  | GTAGTCGGAGCAACGCTAAAGCTAGTCGGGTGGCTTGACCGCGATGGTAAAGGATCTG  | 4020  |
| Dd | 64926 | GTAGTCGGAGCAACGCTAAAGCTAGTCGGGTGGCTTGACCGCGATGGTAAAGGATCTG  | 64985 |
| Qy | 4021  | GCCAGGATCAGACCTTTAATATCTCCGGTTCGGAACAAATAGCACACACAAATTTCT   | 4080  |
| Dd | 64986 | GCCAGGATCAGACCTTTAATATCTCCGGTTCGGAACAAATAGCACACACAAATTTCT   | 65045 |
| Qy | 4081  | ATCTATCGGTGGATAAATCAGGTTAGTGTCTCATATCCATAACCGCTTCAGCAGAGC   | 4140  |
| Dd | 65046 | ATCTATCGGTGGATAAATCAGGTTAGTGTCTCATATCCATAACCGCTTCAGCAGAGC   | 65105 |
| Qy | 4141  | CGCGAGCTTCTGTTTAAAAATCGAGAAAGATTTAAAAATTCAGCCATGAGTCGTAAG   | 4200  |
| Dd | 65106 | CGCGAGCTTCTGTTTAAAAATCGAGAAAGATTTAAAAATTCAGCCATGAGTCGTAAG   | 65165 |
| Qy | 4201  | TAGCAAGTTTGTGGAGGGTCTTCTAACACTTCTGGTGACATTTTCTCCCAACTAT     | 4260  |
| Dd | 65166 | TAGCAAGTTTGTGGAGGGTCTTCTAACACTTCTGGTGACATTTTCTCCCAACTAT     | 65225 |
| Qy | 4261  | TGTTTTAAAGCCAACTTCCATAAATCCCGCTCTGGTAATTAATTAATTAATTAATTA   | 4320  |
| Dd | 65226 | TGTTTTAAAGCCAACTTCCATAAATCCCGCTCTGGTAATTAATTAATTAATTAATTA   | 65285 |
| Qy | 4321  | ACTTTTGTACTTCTGTGAGAAATCCACAGAGGTGGCGCGTATAAGCAGTTCCGCTT    | 4380  |
| Dd | 65286 | ACTTTTGTACTTCTGTGAGAAATCCACAGAGGTGGCGCGTATAAGCAGTTCCGCTT    | 65345 |
| Qy | 4381  | TCACCTACCGTTTATGTGAAAGTCTTCAAACTCGCTTAAACCTCATCAAGCTATATTC  | 4440  |
| Dd | 65346 | TCACCTACCGTTTATGTGAAAGTCTTCAAACTCGCTTAAACCTCATCAAGCTATATTC  | 65405 |
| Qy | 4441  | ATATAATTCACAAATGGAAGCTCTTCTCGGGACACTGTAATTTAACTTTGTAACAAATC | 4500  |
| Dd | 65406 | ATATAATTCACAAATGGAAGCTCTTCTCGGGACACTGTAATTTAACTTTGTAACAAATC | 65465 |
| Qy | 4501  | CTTTATAAAACCGCTTAAGGTGAATCCATTTTACATATTTCTGTTGAGAGCGATCA    | 4560  |
| Dd | 65466 | CTTTATAAAACCGCTTAAGGTGAATCCATTTTACATATTTCTGTTGAGAGCGATCA    | 65525 |
| Qy | 4561  | GCITTTATCGCAATGTTGCTGTACACAGCCCTGGACATAAAGCTAAATTTGTTAAATGA | 4620  |
| Dd | 65526 | GCITTTATCGCAATGTTGCTGTACACAGCCCTGGACATAAAGCTAAATTTGTTAAATGA | 65585 |
| Qy | 4621  | AAAGAAGCAATCTCAACCGCTCGGAAATATAAAGTCACTAAGTGGTATTTTACAA     | 4680  |
| Dd | 65586 | AAAGAAGCAATCTCAACCGCTCGGAAATATAAAGTCACTAAGTGGTATTTTACAA     | 65645 |
| Qy | 4681  | CATGGATTTACTACAAGCCCGGTACAACTGTAATTTAAAAAAAATCCACATGTAATAAT | 4740  |
| Dd | 65646 | CATGGATTTACTACAAGCCCGGTACAACTGTAATTTAAAAAAAATCCACATGTAATAAT | 65705 |
| Qy | 4741  | AAACTTAAGGTATACATCTTCTCAATGTTTATACATAATATGAAGCTTAATCTGTACT  | 4800  |
| Dd | 65706 | AAACTTAAGGTATACATCTTCTCAATGTTTATACATAATATGAAGCTTAATCTGTACT  | 65765 |

|    |       |   |       |
|----|-------|---|-------|
| Qy | 4801  | TATATAATATGTATAGAAATTAATGCCAATAACAGTCTATGTAAGTATATTTTAAAGCATAAT | 4860  |
| Dd | 65766 | TATATAATATGTATAGAAATTAATGCCAATAACAGTCTATGTAAGTATATTTTAAAGCATAAT | 65825 |
| Qy | 4861  | GCAAACTCAAAATAATATAACACAAATTAACCAATTAACCAATTAACCAATTAACCAATTAAC | 4920  |
| Dd | 65826 | GCAAACTCAAAATAATATAACACAAATTAACCAATTAACCAATTAACCAATTAACCAATTAAC | 65885 |
| Qy | 4921  | CCGAAAATACAGAAATGACTACACAAACACACCTGAAACCAATTTTATTTCTCAACATATG   | 4980  |
| Dd | 65886 | CCGAAAATACAGAAATGACTACACAAACACACCTGAAACCAATTTTATTTCTCAACATATG   | 65945 |
| Qy | 4981  | CAACAAATTTTCAGGTAACCACTGTTTAAATAATAGCAGCATGACATTTTATAGCTA       | 5040  |
| Dd | 65946 | CAACAAATTTTCAGGTAACCACTGTTTAAATAATAGCAGCATGACATTTTATAGCTA       | 66005 |
| Qy | 5041  | AGCTCTCAAAAGACAAATTTCTTCAATGATGATCATCAACCTCGGTTTCGTTGTTA        | 5100  |
| Dd | 66006 | AGCTCTCAAAAGACAAATTTCTTCAATGATGATCATCAACCTCGGTTTCGTTGTTA        | 66065 |
| Qy | 5101  | ATGTCGGAATATATAGTCCAGGATACCTTCGTCATCGCCACAAATTAATTTCCAGGGCT     | 5160  |
| Dd | 66066 | ATGTCGGAATATATAGTCCAGGATACCTTCGTCATCGCCACAAATTAATTTCCAGGGCT     | 66125 |
| Qy | 5161  | TGTAATAATCATCCAGTGAATCAGTCTGCTAATCTCAAGATGTTGTTTCTTAATTTTAC     | 5220  |
| Dd | 66126 | TGTAATAATCATCCAGTGAATCAGTCTGCTAATCTCAAGATGTTGTTTCTTAATTTTAC     | 66185 |
| Qy | 5221  | CCATCCAAATCGTGGTGGTGACACTTGGGCAAGACCCCGTAAATTTGCGACGCTCC        | 5280  |
| Dd | 66186 | CCATCCAAATCGTGGTGGTGACACTTGGGCAAGACCCCGTAAATTTGCGACGCTCC        | 66245 |
| Qy | 5281  | ACACAAACGGAGGATGATTTGATCGTTTACGAGGGGACAGATATTTTATATAACA         | 5340  |
| Dd | 66246 | ACACAAACGGAGGATGATTTGATCGTTTACGAGGGGACAGATATTTTATATAACA         | 66305 |
| Qy | 5341  | TCAGAAACGGACCGGTCAATAGCGCACATATGTCATTAAAGATTGGATCGTGACACTG      | 5400  |
| Dd | 66306 | TCAGAAACGGACCGGTCAATAGCGCACATATGTCATTAAAGATTGGATCGTGACACTG      | 66365 |
| Qy | 5401  | GTTGAAATGACGCTGAAAAAATGTTTCTGTCACATGCGTAACTATTTTAAACAC          | 5460  |
| Dd | 66366 | GTTGAAATGACGCTGAAAAAATGTTTCTGTCACATGCGTAACTATTTTAAACAC          | 66425 |
| Qy | 5461  | ACGAGGCGGGAATAAAACGGGTCCCGCGCTAAGCGCTGAATTCGCGACGCGGCTC         | 5520  |
| Dd | 66426 | ACGAGGCGGGAATAAAACGGGTCCCGCGCTAAGCGCTGAATTCGCGACGCGGCTC         | 66485 |
| Qy | 5521  | ATTGCGCAATTCAGATTTGTTAAATCTGTTCCGAACATACAGTCTTAAATCAAGTATAAC    | 5580  |
| Dd | 66486 | ATTGCGCAATTCAGATTTGTTAAATCTGTTCCGAACATACAGTCTTAAATCAAGTATAAC    | 66545 |
| Qy | 5581  | CACGACAAAGTTCAAAACGAGGCAACCGGCTGTTTCCACCGGACATCCCGCTCACA        | 5640  |
| Dd | 66546 | CACGACAAAGTTCAAAACGAGGCAACCGGCTGTTTCCACCGGACATCCCGCTCACA        | 66605 |
| Qy | 5641  | AAACCGCAATCTCTAGGAACGACATCAAGGTTTCAGACAAATTTCCGAACCCGAAACC      | 5700  |
| Dd | 66606 | AAACCGCAATCTCTAGGAACGACATCAAGGTTTCAGACAAATTTCCGAACCCGAAACC      | 66665 |
| Qy | 5701  | CTAGTTCTGAGCACTCTGTTTAAACTGCGGTTAAAGAGACGAGCGGCAACACTGC         | 5760  |
| Dd | 66666 | CTAGTTCTGAGCACTCTGTTTAAACTGCGGTTAAAGAGACGAGCGGCAACACTGC         | 66725 |
| Qy | 5761  | TGCTGTGAGGATTCAGCTTAGTGGCTGACGGCTTCTGACAGGCGGCTAGTCCCTG         | 5820  |
| Dd | 66726 | TGCTGTGAGGATTCAGCTTAGTGGCTGACGGCTTCTGACAGGCGGCTAGTCCCTG         | 66785 |
| Qy | 5821  | GGGGCGCTCGGCACCTTTTGAAGCGGCAACCGCTTCCAGAGCGAGGCAAGCGCTCT        | 5880  |
| Dd | 66786 | GGGGCGCTCGGCACCTTTTGAAGCGGCAACCGCTTCCAGAGCGAGGCAAGCGCTCT        | 66845 |

5981 TTGAGTTCGGCCATATCTCGGCATAGAGATTGCTCGCTGATCTTTTAGTGCTATA 5940  
5982 TTGAGTTCGGCCATATCTCGGCATAGAGATTGCTCGCTGATCTTTTAGTGCTATA 5905  
5941 TATATCTGTGGGTAAACAACTCACTCCCTGTAAAGGGGAATAAGGTCGCTCTTTCT 6000  
5996 TATATCTGTGGGTAAACAACTCACTCCCTGTAAAGGGGAATAAGGTCGCTCTTTCT 6695  
6001 ATCGTTTCTCTGAACTATAGTTACATCTTTAAGTATAACCAAGAGCGCTAAACCA 6060  
6966 ATCGTTTCTCTGAACTATAGTTACATCTTTAAGTATAACCAAGAGCGCTAAACCA 67025  
6061 GGATGGTCGACACCTTTAAAAATCGCAAGCGACCGGATGGCTCTATPCGCGTTTCCA 6120  
67026 GGATGGTCGACACCTTTAAAAATCGCAAGCGACCGGATGGCTCTATPCGCGTTTCCA 67085  
6121 TGTCTCTTCAAGAGTAGTACGATCTCCAGTGTAGACTCTGTCTCCCTGTTAACACC 6180  
67086 TGTCTCTTCAAGAGTAGTACGATCTCCAGTGTAGACTCTGTCTCCCTGTTAACACC 67145  
6181 TTAACGTGAATGAAACAGCAGTACGACCGTGAACCTTCCTGATGTCTCTTAAACCTAAC 6240  
67146 TTAACGTGAATGAAACAGCAGTACGACCGTGAACCTTCCTGATGTCTCTTAAACCTAAC 67205  
6241 CCGAAACAGAGCTAAATACCAATGACTGTACCCCTACCCGCAAGCAAGCCCGCTACTA 6300  
67206 CCGAAACAGAGCTAAATACCAATGACTGTACCCCTACCCGCAAGCAAGCCCGCTACTA 67265  
6301 TTAGACAGAGGTGAGTAAACCAAGCTATCTTTTAAACCCATACGTTGAGTTTGAAGT 6360  
67266 TTAGACAGAGGTGAGTAAACCAAGCTATCTTTTAAACCCATACGTTGAGTTTGAAGT 67325  
6361 AAACAGCTGTGTATATTTTACAGCGCTGTATATAAATGGATACCGACGACATCAGTGA 6420  
67326 AAACAGCTGTGTATATTTTACAGCGCTGTATATAAATGGATACCGACGACATCAGTGA 67385  
6421 ATTAACTTTTATTTTACAGTTTCAAGATTCAGGAGGGCGTGTACAGTGTGTAGTGGAGCA 6480  
67386 ATTAACTTTTATTTTACAGTTTCAAGATTCAGGAGGGCGTGTACAGTGTGTAGTGGAGCA 67445  
6481 TATCTCGCCAAATGGGAATGGCTGATGAATCCACACTTAGTGTCTCGATCAGTGTCTGT 6540  
67446 TATCTCGCCAAATGGGAATGGCTGATGAATCCACACTTAGTGTCTCGATCAGTGTCTGT 67505  
6541 GTACTTTACTAGTCCGATGAACCTAAATTTGGCGTGTATAGTGTGTACCAACCAAAAA 6600  
67506 GTACTTTACTAGTCCGATGAACCTAAATTTGGCGTGTATAGTGTGTACCAACCAAAAA 67565  
6601 CTGCTGGCCTCGCTTTTAAAGCTCTATGCAAGCTTTAATCGTGTCCAATAGCTTGT 6660  
67566 CTGCTGGCCTCGCTTTTAAAGCTCTATGCAAGCTTTAATCGTGTCCAATAGCTTGT 67625  
6661 TTTAAATGCAACTCTGGTACATCTTTATGACAGTGTGTCCAAAAAACAAGATTAA 6720  
67626 TTTAAATGCAACTCTGGTACATCTTTATGACAGTGTGTCCAAAAAACAAGATTAA 67685  
6721 AAACACCTTAACTCGCTGTCTGTGTAGTCTCTGTATATAACCTCTCAACCAAGAAAAA 6780  
67686 AAACACCTTAACTCGCTGTCTGTGTAGTCTCTGTATATAACCTCTCAACCAAGAAAAA 67745  
6781 TTTTAAACCAAACTCGCCAGGTACTGAAACGATGTGACGGAACAGTGTGAAAAGTGT 6840  
67746 TTTTAAACCAAACTCGCCAGGTACTGAAACGATGTGACGGAACAGTGTGAAAAGTGT 67805  
6841 TATCATCTCTTCACTCCCTTTTGAAGTTTGTGTACACCGACCATGTGAGATAGGCA 6900  
67806 TATCATCTCTTCACTCCCTTTTGAAGTTTGTGTACACCGACCATGTGAGATAGGCA 67865  
6901 CCAGCTCAGGGGGAGGTGCGATCTGAGGGGTAAACAATTCGTGGCGTGGGGGTACTC 6960  
67866 CCAGCTCAGGGGGAGGTGCGATCTGAGGGGTAAACAATTCGTGGCGTGGGGGTACTC 67925  
6961 CAGTTTATAGATCTCCGCGAGGCGGTACGAGTTCTTCTGCTCAATCTACGACGCCGC 7020

67926 CAGTTTATAGATCTCCGCGAGGCGGTACGAGTTCTTCGCTCAATCTACGACGCCGC 67985  
7021 ATCCACAGGGGAGTACTGAGTTAAATAGTCCCGGAATCTATCGTTCGTAATTTTGTG 7080  
67986 ATCCACAGGGGAGTACTGAGTTAAATAGTCCCGGAATCTATCGTTCGTAATTTTGTG 68045  
7081 GAGCTCTCTTTAGACAGTAAACCACTTGTCTCCGACACTCGGTACGATGAACCGCTCTCC 7140  
68046 GAGCTCTCTTTAGACAGTAAACCACTTGTCTCCGACACTCGGTACGATGAACCGCTCTCC 68105  
7141 TTGTCTGTGTGATAGACACCTGGAGTATCTTAGACAGGTATAGGGCGGTTCCACTCT 7200  
68106 TTGTCTGTGTGATAGACACCTGGAGTATCTTAGACAGGTATAGGGCGGTTCCACTCT 68165  
7201 AAGGTTTCCGCTTGTGTAGAGCGGATTCGCGGCAATTCAGCGTCAAGTGTGTTCATTCCAC 7260  
68166 AAGGTTTCCGCTTGTGTAGAGCGGATTCGCGGCAATTCAGCGTCAAGTGTGTTCATTCCAC 68225  
7261 GACTATCATGCGAGTGGGGATAGTGAACCTTGTATCATATGATTTAAAAGATGGTCTTT 7320  
68226 GACTATCATGCGAGTGGGGATAGTGAACCTTGTATCATATGATTTAAAAGATGGTCTTT 68285  
7321 CCGCTTAATGGCGACTCATGTCTGTAAACGTTTACAGGTTTAAACCACTCAAGAG 7380  
68286 CCGCTTAATGGCGACTCATGTCTGTAAACGTTTACAGGTTTAAACCACTCAAGAG 68345  
7381 GTCTGTGCGTGAACCTTTTGTGCGGACTCTCCGAAACCGCTCAGGATAGTGTACACCGC 7440  
68346 GTCTGTGCGTGAACCTTTTGTGCGGACTCTCCGAAACCGCTCAGGATAGTGTACACCGC 68405  
7441 GGTCTCTTATTAACCAAGCATACGTCAAAAAACATGACCCAGCTCACTTAAATAAGCT 7500  
68406 GGTCTCTTATTAACCAAGCATACGTCAAAAAACATGACCCAGCTCACTTAAATAAGCT 68465  
7501 CACAGAGAACTGTCTCGATGTATTTGAATGGAATGTCCTCCCTGGGTGCCATCTCCGG 7560  
68466 CACAGAGAACTGTCTCGATGTATTTGAATGGAATGTCCTCCCTGGGTGCCATCTCCGG 68525  
7561 GCTCATCGCGATCTGATCTCTTTTAACTGTCTGCTCTACCGGGTTCCTCCGGGTAA 7620  
68526 GCTCATCGCGATCTGATCTCTTTTAACTGTCTGCTCTACCGGGTTCCTCCGGGTAA 68585  
7621 AACCCGCGGGGCCACCTGTAAACGTCCTCGTGGCAGAAATGCGCCAGGCACTGTGAG 7680  
68586 AACCCGCGGGGCCACCTGTAAACGTCCTCGTGGCAGAAATGCGCCAGGCACTGTGAG 68645  
7681 AATCTCACCGAGAGGGCCCTTTGTGTACCGAAAGATGTTTATAGCTCTGCTGCTGAG 7740  
68646 AATCTCACCGAGAGGGCCCTTTGTGTACCGAAAGATGTTTATAGCTCTGCTGCTGAG 68705  
7741 CGGGTCTGTTATACCGCCCACTGTGGCCAGAGTTTCTGCACGACGTCTACGCGAAATGAA 7800  
68706 CGGGTCTGTTATACCGCCCACTGTGGCCAGAGTTTCTGCACGACGTCTACGCGAAATGAA 68765  
7801 GGCCAAAGTGTGGGGCGGTGCTCTCATATGTTGAGGGCGCTATCATGGCTAT 7860  
68766 GGCCAAAGTGTGGGGCGGTGCTCTCATATGTTGAGGGCGCTATCATGGCTAT 68825  
7861 TGCCGACTCGGTCTCTGCTCACTTATATACCTCGATGCGCGGAAACCTGGAACCTTAG 7920  
68826 TGCCGACTCGGTCTCTGCTCACTTATATACCTCGATGCGCGGAAACCTGGAACCTTAG 68885  
7921 ACTCAAGCACTGTGCAAACTTGTGTTCATCCCATCTTCTTCAAGAAATCTTAGCCCC 7980  
68886 ACTCAAGCACTGTGCAAACTTGTGTTCATCCCATCTTCTTCAAGAAATCTTAGCCCC 68945  
7981 TATCAGCTACTGCGCAACCGGAAGATGGTTCCTGACAAATATTTTACCATCACCGGTAC 8040  
68946 TATCAGCTACTGCGCAACCGGAAGATGGTTCCTGACAAATATTTTACCATCACCGGTAC 69005  
8041 GCGCGAGAGAGGGCCCTCTGCTGACTGGAAGTACTAGCGGATGAGCTGTCCGGAGAG 8100

69006 GCGCGAGAGAGCGCCCTGTCGTGACTTGAAGTACTAGCGGATGACGCTGTCGCGGAG 69065  
QY 8101 CAGCTTGTCCCGATTCCTTAATCTCTGCGAGTATCGGAACCGGGCTTCTCCGCGACC 8160  
Db 69066 CAGCTTGTCCCGATTCCTTAATCTCTGCGAGTATCGGAACCGGGCTTCTCCGCGACC 69125  
QY 8161 CTTGGTTGACCTCAGTAATGTTCTTAGAAATCAGAAATCATCTCAGCGCCGCCACCCCT 8220  
Db 69126 CTTGGTTGACCTCAGTAATGTTCTTAGAAATCAGAAATCATCTCAGCGCCGCCACCCCT 69185  
QY 8221 GAGTCATTTGTCATCACAACACGACGCCAGTTCCTCAGTACGTACGATCAGCATATTATAC 8280  
Db 69186 GAGTCATTTGTCATCACAACACGACGCCAGTTCCTCAGTACGTACGATCAGCATATTATAC 69245  
QY 8281 GCCAACCCAGGGGTTGTTCCCGCCCAATGTTTATGACACAGTGAAGACGGTGTGACA 8340  
Db 69246 GCCAACCCAGGGGTTGTTCCCGCCCAATGTTTATGACACAGTGAAGACGGTGTGACA 69305  
QY 8341 GAGCATTTACACAGGACAGACGGCTATTTTGGCCGCGCACTAACCGGTTGACATC 8400  
Db 69306 GAGCATTTACACAGGACAGACGGCTATTTTGGCCGCGCACTAACCGGTTGACATC 69365  
QY 8401 TCGGGCCCTGCGCCCATATCGCATGTTCCCGAGTTCGGGACGCTCTCGCAGGTGA 8460  
Db 69366 TCGGGCCCTGCGCCCATATCGCATGTTCCCGAGTTCGGGACGCTCTCGCAGGTGA 69425  
QY 8461 AGGCTCGCGGGCGTTCGATTGCGGGAACACGATCGCTTCGCGACCGCGTTACCGCC 8520  
Db 69426 AGGCTCGCGGGCGTTCGATTGCGGGAACACGATCGCTTCGCGACCGCGTTACCGCC 69485  
QY 8521 CAATGTCGCGCAAGAGATGAAACGGTAGCACAGTTGGGAACGCTCCCGTAAAAA 8580  
Db 69486 CAATGTCGCGCAAGAGATGAAACGGTAGCACAGTTGGGAACGCTCCCGTAAAAA 69545  
QY 8581 CGTCCATCGGAGGCGGTATACGCTCCTCCTGTTTATATACCAATAGACTTAC 8640  
Db 69546 CGTCCATCGGAGGCGGTATACGCTCCTCCTGTTTATATACCAATAGACTTAC 69605  
QY 8641 GTCCCCGTAGGTCGCGCCAGAGTCGCGCGATATCGCAACACTCCAGATCCCGCAT 8700  
Db 69606 GTCCCCGTAGGTCGCGCCAGAGTCGCGCGATATCGCAACACTCCAGATCCCGCAT 69665  
QY 8701 GCGGGCGGCTCTCCGCCCTTCGCGCAACCGCGCAACGGTCCCGCTTAAGAGAAGA 8760  
Db 69666 GCGGGCGGCTCTCCGCCCTTCGCGCAACCGCGCAACGGTCCCGCTTAAGAGAAGA 69725  
QY 8761 GCCACGCGAGAGCTGGCAGACAAAAGACTGAAGGCGGACGTTCCGGGCGCGCAACAGT 8820  
Db 69726 GCCACGCGAGAGCTGGCAGACAAAAGACTGAAGGCGGACGTTCCGGGCGCGCAACAGT 69785  
QY 8821 AAACCCACCTTTCCCGGACCGTCCGCGATGCGCGTTCGCGAGCAGGGCTTATTCGATT 8880  
Db 69786 AAACCCACCTTTCCCGGACCGTCCGCGATGCGCGTTCGCGAGCAGGGCTTATTCGATT 69845  
QY 8881 AATCGAAAGCTCCACGATGTAACCGCGACGATCTGACACCAAAACGACGACAT 8940  
Db 69846 AATCGAAAGCTCCACGATGTAACCGCGACGATCTGACACCAAAACGACGACAT 69905  
QY 8941 GCTAGCGGCTATCCTACAGGACCTGTATGACATACAGTCCCCCGCGCATCGATTCCCC 9000  
Db 69906 GCTAGCGGCTATCCTACAGGACCTGTATGACATACAGTCCCCCGCGCATCGATTCCCC 69965  
QY 9001 CTCACGAACCTCGACATGAGAGATATTTCCAGAGGTTAGTCCGCCATCTAGCGGCA 9060  
Db 69966 CTCACGAACCTCGACATGAGAGATATTTCCAGAGGTTAGTCCGCCATCTAGCGGCA 70025  
QY 9061 CGGATCGCCTTGA 9073  
Db 70026 CGGATCGCCTTGA 70038

LOCUS AF210726 130733 bp DNA linear VRL 27-MAR-2000  
DEFINITION Macaca mulatta rhadinovirus 26-95 long unique region L-DNA,  
complete sequence.  
ACCESSION AF210726  
VERSION AF210726.1 GI:7329990  
KEYWORDS  
SOURCE  
ORGANISM  
Macaca mulatta rhadinovirus 26-95  
Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
Gammaherpesvirinae; Rhadinovirus.  
REFERENCE  
1. (bases 1 to 130733)  
Alexander, L., Denekamp, L., Knapp, A., Auerbach, M. R., Damania, B. and  
Desrosiers, R. C.  
The primary sequences of rhesus monkey rhadinovirus isolate 26-95:  
sequence similarities to Kaposi's sarcoma-associated herpesvirus  
and rhesus monkey rhadinovirus isolate 17577  
J. Virol. 74 (7), 3388-3398 (2000)  
20173730  
MEDLINE  
PUBMED  
10708456  
REFERENCE  
2. (bases 1 to 130733)  
Alexander, L., Denekamp, L. M., Knapp, A., Auerbach, M., Czajak, S.,  
Damania, B. and Desrosiers, R. C.  
Direct Submission  
Submitted (02-DEC-1999) Microbiology, New England Regional Primate  
Research Center, One Pinehill Dr, Southborough, MA 01772, USA  
Location/Qualifiers  
1. 130733  
/organism="Macaca mulatta rhadinovirus 26-95"  
/mol\_type="genomic DNA"  
/isolate="Macaca mulatta rhadinovirus isolate 26-95"  
/db\_xref="taxon:119193"  
513..1784  
/gene="R1"  
513..1784  
/gene="R1"  
/codon\_start=1  
/product="ORF1"  
/protein\_id="AAF59980.1"  
/db\_xref="GI:7329991"  
/translation="MFVLVFLMLQPVSVLLPAKLVSPTWCPHPGDTVLLTCRGT  
STARDQSTQFRNNTLMRGSNFYGLVSVTPNATISDRYACQTKTKTRNNIDFRVS  
LRTQTQGRGSSYGYTANNTRVLRCSGNGTVLNVLNVLHNGTAVNGTRNTHFV  
LSTKGTGFCSAFIGNEKGYVSOTINVFTSFPTKPTNDIPNESHFNKTIQIQTASY  
QHPNVYVSPVPSIGVLGIALSLIMLCLFTRCNSESNSTNSVASOTSYIQPSH  
NQRNTNCRSHYRNAHQBSHELPHQHTSETDSCQLVLLKVNVAVDGQENIT  
NEVMEQDVVVVENIETQSDNEVDHNDYSDTINPNFNYSGLILEEDVEVFNLEN  
QYHLILENDHNEYHNLNMLEQYDWLE"  
complement (1852..2418)  
/gene="ORF02"  
complement (1852..2418)  
/gene="ORF02"  
/codon\_start=1  
/product="DHFR"  
/protein\_id="AAF59981.1"  
/db\_xref="GI:7329992"  
/translation="MDIATNCIVAVDQLGKGNKTMPPVYLNRNMYFQWMTSPSV  
VGEKNVIMGRKWTFSIPEKRPVAVNINILSRELPEPHGAHFLATLDDAFNFR  
QYKLKEQLNTVWIGGKSVESVLNKKPLKLYITRIMESFDCDVFPSINFTEYTM  
SELFQKDTNFENGIGIKYKQYIEKNFK"  
2595..3782  
/gene="ORF04"  
2595..3782  
/gene="ORF04"  
/codon\_start=1  
/product="complement binding protein"  
/protein\_id="AAF59982.1"  
/db\_xref="GI:7329993"  
/translation="MFWIPITLCFCQVILVDSKDGENVCLKPPFDRFMVKTAQNE  
NYAVGTRVELTCRPFYQLQAVYVEGLSNGTWTTPNAECRKRKCSPEIDILNGEVII  
TDSNAPKFGSNITKYKNTGYLLGAVRTCLLYKDSNLNDWQDPAAPTCIEKCKQP  
DIENGKYPVQEFYNYLETITFTCNKDPFLIGNTTTCMTNGTWSVPVKCOQITCSA  
PNIDHGTLLVGSSEHYKGGOSVTIGCEDGTLNGHKMCTCEYSLWNPLPTCPINKT  
DPTSEVPSGCTKEQEMPTENPKSHSESTTTTETPKTQTHKSTSPSPKVPNPETHK



Dbb 59154 GGGCGGTTCTTGGCGTTTACCGTTTATCTAAATTACCGGAACTGCTGGTCTCGGAAAAGT 59213  
Qy 241 ACCAGCATATCGCCCTGTACCAAAATTTAAACTGCTGTATCAACCGGCGCGACCAACATA 300  
Db 59214 ACCAGCATATCGCCCTGTACCAAAATTTAAACTGCTGTATCAACCGGCGCGACCAACATA 59273  
Qy 301 GCCGCACAGAACTATCGCGTCCCTAAAGAGGTTCTGTCCACGATCTTTCAGCGCTTTT 360  
Db 59274 GCCGCACAGAACTATCGCGTCCCTAAAGAGGTTCTGTCCACGATCTTTCAGCGCTTTT 59333  
Qy 361 GGCTTTAAGAGCGGACACATCAATATAGCCGTCAGAAAAAGCTCATCAGACCGAGCCGTA 420  
Db 59334 GGCTTTAAGAGCGGACACATCAATATAGCCGTCAGAAAAAGCTCATCAGACCGAGCCGTA 59393  
Qy 421 TCCATAGAGCAATTCAGCAACAGGAGCTATCGAAGTATGGCGGTATAGTGACATTT 480  
Db 59394 TCCATAGAGCAATTCAGCAACAGGAGCTATCGAAGTATGGCGGTATAGTGACATTT 59453  
Qy 481 ATGAAAGAGGTTATGGCGGAAAAAACCCTATGCGATATGCGGACTATATCCAAACCGAAT 540  
Db 59454 ATGAAAGAGGTTATGGCGGAAAAAACCCTATGCGATATGCGGACTATATCCAAACCGAAT 59513  
Qy 541 TTTGAAACCTCTCGAGAGATGACCGGACCGGTGTTATGACCTTCCAAATATATGTAATC 600  
Db 59514 TTTGAAACCTCTCGAGAGATGACCGGACCGGTGTTATGACCTTCCAAATATATGTAATC 59573  
Qy 601 GACGAGCGCGGAAACCTGCTCTTACATACCTACACACCGTCTGTCTTTTACTGGTTC 660  
Db 59574 GACGAGCGCGGAAACCTGCTCTTACATACCTACACACCGTCTGTCTTTTACTGGTTC 59633  
Qy 661 CTAAACAGCTGGCTAAATACCCCTCTTTACCGCACCGGCGCGGTTCGCGCATATATGC 720  
Db 59634 CTAAACAGCTGGCTAAATACCCCTCTTTACCGCACCGGCGCGGTTCGCGCATATATGC 59693  
Qy 721 GTGCGTTCGCCAACGACAGAAACCGGTTCCAGTCAACTTACCAACCGGACGCAAAAG 780  
Db 59694 GTGCGTTCGCCAACGACAGAAACCGGTTCCAGTCAACTTACCAACCGGACGCAAAAG 59753  
Qy 781 ACCGAGATATCGTGTGCGAGAACATCCTAAACATTCATGATAGGGAAGAGTCTGTATCT 840  
Db 59754 ACCGAGATATCGTGTGCGAGAACATCCTAAACATTCATGATAGGGAAGAGTCTGTATCT 59813  
Qy 841 GAGTAGTACATCTGAGAGAAACTGGGCGTGTGTTATAAAACAAAGCGCTGCACCGAT 900  
Db 59814 GAGTAGTACATCTGAGAGAAACTGGGCGTGTGTTATAAAACAAAGCGCTGCACCGAT 59873  
Qy 901 CTGCACTTCGGACACCTCTTAAAAATTTTAGAGTATAATCTTCCCATTCCTGACCAAGTC 960  
Db 59874 CTGCACTTCGGACACCTCTTAAAAATTTTAGAGTATAATCTTCCCATTCCTGACCAAGTC 59933  
Qy 961 ATGAGTTAGTGACAGATTTGTGCTTCTTAAAGTAAGATATGATCTCTTTTAAATATAC 1020  
Db 59934 ATGAGTTAGTGACAGATTTGTGCTTCTTAAAGTAAGATATGATCTCTTTTAAATATAC 59993  
Qy 1021 ATTGGCTGGACCGACTCTTTTGTCCATATAGGAGTAAAGCGGTATCTTAAACACCTA 1080  
Db 59994 ATTGGCTGGACCGACTCTTTTGTCCATATAGGAGTAAAGCGGTATCTTAAACACCTA 60053  
Qy 1081 CACACATGTCTAACGCTAGGGGCGATACAGGACACGAAGCTCTTTTACCTGTCCCGTG 1140  
Db 60054 CACACATGTCTAACGCTAGGGGCGATACAGGACACGAAGCTCTTTTACCTGTCCCGTG 60113  
Qy 1141 GTGTGCGAGGTTTGTGAGCGGTTTGAAGATACAGCGGCGGCTCAACCTCCCAAC 1200  
Db 60114 GTGTGCGAGGTTTGTGAGCGGTTTGAAGATACAGCGGCGGCTCAACCTCCCAAC 60173  
Qy 1201 CTCACCGTACCGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATTTATTCGAGTTT 1260  
Db 60174 CTCACCGTACCGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATTTATTCGAGTTT 60233  
Qy 1261 GTGGACAGGACATGTCATAGTCGCCACGGAATCCATGAAAGCTTGCACCGAGTTACC 1320  
Db 60234 GTGGACAGGACATGTCATAGTCGCCACGGAATCCATGAAAGCTTGCACCGAGTTACC 60293

Qy 1321 TTTATCAACCAAGTTTGTAAACACAGCCACGTCGTCTTTAAATGAAAAACAAAAATGC 1380  
Db 60294 TTTATCAACCAAGTTTGTAAACACAGCCACGTCGTCTTTAAATGAAAAACAAAAATGC 60353  
Qy 1381 ATATGTGGTTTCAAGGTACGTACTTTTGTAGTTCAAAAGAAATCCCTAGACAGCAACTCTTC 1440  
Db 60354 ATATGTGGTTTCAAGGTACGTACTTTTGTAGTTCAAAAGAAATCCCTAGACAGCAACTCTTC 60413  
Qy 1441 GTGAGAGCGCATTCGCAAGACCGTCCCGAATAACGTATACCGTCTTTTAAACACAGCTACTG 1500  
Db 60414 GTGAGAGCGCATTCGCAAGACCGTCCCGAATAACGTATACCGTCTTTTAAACACAGCTACTG 60473  
Qy 1501 TACAAACCGCATGTACTCGTTTTCACCGCTACGGGTGACTAGGTACACAGCAAAATATCTG 1560  
Db 60474 TACAAACCGCATGTACTCGTTTTCACCGCTACGGGTGACTAGGTACACAGCAAAATATCTG 60533  
Qy 1561 CAAGACCTTAAATTTGACACCCCTCCCGCGCGCTCTGCAACACCGGCGGTGTAGACCTTCAA 1620  
Db 60534 CAAGACCTTAAATTTGACACCCCTCCCGCGCGCTCTGCAACACCGGCGGTGTAGACCTTCAA 60593  
Qy 1621 ACGETTCGTGAAGAGTTAAACCTGGAAGACAGACATCTTACACAGTGTGAGTCCCGCG 1680  
Db 60594 ACGETTCGTGAAGAGTTAAACCTGGAAGACAGACATCTTACACAGTGTGAGTCCCGCG 60653  
Qy 1681 CCACCCCGGCTATCACTCCCTCCAGTCTTGGTCCACACGTACTGCGCGCTTAAAGGAC 1740  
Db 60654 CCACCCCGGCTATCACTCCCTCCAGTCTTGGTCCACACGTACTGCGCGCTTAAAGGAC 60713  
Qy 1741 GTGTTTCGCTCCAGAAATAAAGTGGCTGTGCTGTTTGGCGGGAGTTTGAAGAGAA 1800  
Db 60714 GTGTTTCGCTCCAGAAATAAAGTGGCTGTGCTGTTTGGCGGGAGTTTGAAGAGAA 60773  
Qy 1801 ACGTTTTCGCGTTTACGGTTTACATGCTGAAGGACCGAGTTGACTTTGTCTCCCT 1860  
Db 60774 ACGTTTTCGCGTTTACGGTTTACATGCTGAAGGACCGAGTTGACTTTGTCTCCCT 60833  
Qy 1861 TCAGAACGTCTCAACCGGCTGTTGGCGCTTTCGATCGACCGTGAATCGTATAAAATTAAG 1920  
Db 60834 TCAGAACGTCTCAACCGGCTGTTGGCGCTTTCGATCGACCGTGAATCGTATAAAATTAAG 60893  
Qy 1921 GGGTACACGTTTTTACCGGTAGCGTTCGCTGCTCAGGGTTTGCACCTCAGTGATGAC 1980  
Db 60894 GGGTACACGTTTTTACCGGTAGCGTTCGCTGCTCAGGGTTTGCACCTCAGTGATGAC 60953  
Qy 1981 CTCAGGAAGAGATGCGCTCCCTGCTGCTGAGGACTCTAGCGGTTTTTATCGCGTGCTA 2040  
Db 60954 CTCAGGAAGAGATGCGCTCCCTGCTGCTGAGGACTCTAGCGGTTTTTATCGCGTGCTA 61013  
Qy 2041 GAGAAATAACATAACCAATTTGACCGAAACCATGAGGACCGGAGCGTTTTCCAGTGTGC 2100  
Db 61014 GAGAAATAACATAACCAATTTGACCGAAACCATGAGGACCGGAGCGTTTTCCAGTGTGC 61073  
Qy 2101 TGTGCGGGGACTATGGGTCAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGA 2160  
Db 61074 TGTGCGGGGACTATGGGTCAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGA 61133  
Qy 2161 ATGTCGTTGAGCGAGTAGCGGTAGTATTTGGTCCCAAGAACGCTCCAGACAGCCAC 2220  
Db 61134 ATGTCGTTGAGCGAGTAGCGGTAGTATTTGGTCCCAAGAACGCTCCAGACAGCCAC 61193  
Qy 2221 GTGTATGTAGCAATATCAAGGGCTGTTAACTCAAAATTTTGTCTAGGACAGCAACCCC 2280  
Db 61194 GTGTATGTAGCAATATCAAGGGCTGTTAACTCAAAATTTTGTCTAGGACAGCAACCCC 61253  
Qy 2281 CTTAAAAACCTCTCTCAGAGAACCGATCGATAACACCTCCCGCAAGCATATAGTCCGCGC 2340  
Db 61254 CTTAAAAACCTCTCTCAGAGAACCGATCGATAACACCTCCCGCAAGCATATAGTCCGCGC 61313  
Qy 2341 CTCACACCCCAACCAACACCTCTACTTAATTAATAAACAAGCAAAATGGAAC 2400  
Db 61314 CTCACACCCCAACCAACACCTCTACTTAATTAATAAACAAGCAAAATGGAAC 61373





|    |       |   |       |
|----|-------|---|-------|
| Db | 63525 | GTTCCAACTTTATATCGCAAGGTTGCAACGCACCGTCTTGCGCACCATGCGCAAACTCTGTA  | 63584 |
| Qy | 4615  | AAATGAAAAAGAGCAATCTCAAAACGGCTCCGGAATATAAAAGTCACTAAAGTGGTATATT   | 4674  |
| Db | 63585 | AACTGGAAGAAACAATTTCAAATGGTGCAGGATATTAACTCCAGTTAGGGGACAAATA      | 63644 |
| Qy | 4675  | TTACACATCGGATTTACTACAAGCCGGTACACATGTAAATTAABAAAAATCCACATGTA     | 4734  |
| Db | 63645 | TTACACATCGGATTTGGAAATTTAAACCAACGAGTACCCAAATATATAGTATAAAAATAAAAT | 63704 |
| Qy | 4735  | ATAAATAAACTTAAGGTATACATACCTCTCATCTGTTTTATACATAATATGAACGCTAATGC  | 4794  |
| Db | 63705 | GTAAAAAACTTAATGTAACATATTTCTGTATTCTTACTTAATGTAAATCTGNACTT        | 63764 |
| Qy | 4795  | TGTACTTATATATATGTATAGNAATTTATCCCAATACAGTCATGGTAAAGTATATTTAAGG   | 4854  |
| Db | 63765 | CGTACTTATATATACGTTGTAAACT-----GCATGTATTTGTATPAA                 | 63806 |
| Qy | 4855  | CATAATGCAAAACATCAAAATAATATAACAAAAATGCAAGCTTCCGGATATCGGCCAGCT    | 4914  |
| Db | 63807 | GTTAATAGATTAT-AAAAGAAAGTGCTAAAAATAGTACTTCCGGATATCGGCCAGCT       | 63865 |
| Qy | 4915  | GTTAACCCGAAATAACAGAAATGACTACACAAACACACCTGAAACCAATTTTATTTCTCAA   | 4974  |
| Db | 63866 | GTTAACCCGAAATAACAGAAATGACTACACAAACACACCTGAAACCAATTTTATTTCTCAA   | 63925 |
| Qy | 4975  | CATATGCAACAATTTTACGGGTAAACACCATGTTTAATAAATAACGACGATGCACATTTT    | 5034  |
| Db | 63926 | CACATGCAACAATTTTACAGTACACCAATGTTTAATAAATAACGACGATGCACATTTT      | 63985 |
| Qy | 5035  | TAGCTAAGCTCTTCAAAGACAAATTTCTCTTCAATGTATGACTCATCAACTCGGTTTCC     | 5094  |
| Db | 63986 | TAGCTAAGCTCTTCAAAGACAAATTTCTCTTCAATGTATGACTCATCAACTCGGTTTCC     | 64045 |
| Qy | 5095  | GTGTTAATGTCGGAATATATAGTCCAGGATACCTCGTCATCGCCACAAATTAATTC        | 5154  |
| Db | 64046 | GTGTTAATGTCGGAATATATAGTCCAGGATACCTCGTCATCGCCACAAATTAATTC        | 64105 |
| Qy | 5155  | AGGGCTGTAAAAATATATCCAGTGAATCAGCTGCTTAATGTCAAGGATGTTGTTTCTAAT    | 5214  |
| Db | 64106 | AGGGCTGTAAAAATATATATCCAGTGAATCAGCTGCTTAATGTCAAGGATGTTGTTTCTAAT  | 64165 |
| Qy | 5215  | TTTAGCCATCCAAATGGTGAGTGGGTGACATTTTGGGACAGACCCCGTAAAAATGGCA      | 5274  |
| Db | 64166 | TTTAGCCATCCAAATGGTGAGTGGGCGACATTTTGGGACAGACCCCGTAAAAATGGCA      | 64225 |
| Qy | 5275  | CGCTCCACACAAACGGAAGATGCAATTCGATCGTTCAACAGGGGAGAGATATTTATTTTT    | 5334  |
| Db | 64226 | CGCTCCACACAAACGGAAGATGCAATTTGATCGTTCAACAGGGGAGAGATATTTATTTTT    | 64285 |
| Qy | 5335  | ATAACATCAGGAACCCGTCATATGACGCACATATGTCATTAAGATTGGAGTCTGT         | 5394  |
| Db | 64286 | ATAACATCAGGAACCCGTCATATGACGCACATATGTCATTAAGATTGGAGTCTGT         | 64345 |
| Qy | 5395  | ACACTGTTGAAATGACGGTGAATAAATAATGATTTGTTGTACATGCGTACTATTTTTT      | 5454  |
| Db | 64346 | ACACTGTTGAAATGACGGTGAATAAATAATGATTTGTTGTACATGCGTAACTATTTTTT     | 64405 |
| Qy | 5455  | AACACACGAGGCGGAATAAAAAACGGGTCCCCGGCCTAAGGCCGTGAATGCGGAGGCA      | 5514  |
| Db | 64406 | AACACACGAGGCGGAATAAAAAACGGGTCCCCGGCCTAAGGCCGTGAATGCGGAGGCA      | 64465 |
| Qy | 5515  | CGGTCATTCGCCATTTCCAGATTTGTTTAATCTGTTTCGAAACATACAGTCTAAAAATCAAG  | 5574  |
| Db | 64466 | CGGTCATTCGCCATTTCCAGATTTGTTTAATCTGTTTCGAAACATACAGTCTAAAAATCAAG  | 64525 |
| Qy | 5575  | TATAACCAACGACAAAGTTTCAACAGGAGGCAACCGCGCTGGTTTCCACCCGGACATCCCG   | 5634  |
| Db | 64526 | TATAACCAACGACAAAGTTTCAACAGGAGGCAACCGCGCTGGTTTCCACCCGGACATCCCG   | 64585 |
| Qy | 5635  | CTCAAAAACCCGAATCTCTTAGAAACGCACATCAAAAGCGTTTCAGACAAAATTCGAAACC   | 5694  |
| Db | 64586 | CTCAAAAACCCGAATCTCTTAGAAACGCACATCAAAAGCGTTTCAGACAAAATTCGAAACC   | 64645 |



QY 6775 AAAAAATTTTAAACCAACTCGCCAGGTACTGAAACGATGTGACGACAGTCTGTGAAA 6834  
 Db 65726 AAAAAATTTTAAACCAACTCGCCAGGTACTGAAACGATGTGACGACAGTCTGTGAAA 65785  
 QY 6835 AGTGTCTATCATCTCTTACCTCCCTTTGAGGTTTGGTTTACACGACCATCTGAGA 6894  
 Db 65786 AGTGTCTATCATCTCTTACCTCCCTTTGAGGTTTGGTTTACACGACCATCTGAGA 65845  
 QY 6895 TAGGCACCAAGCTCAGGGGAGGTGCGATCGTAGGGGTAAACAATTCGTGGCGGTGGG 6954  
 Db 65846 TAGGCACCAAGCTCAGGGGAGGTGCGATCGTAGGGGTAAACAATTCGTGGCGGTGGG 65905  
 QY 6955 GTACTCAGTTTATAGATCTCCGACGAGGACGCTGACGAGTTCTTGGTGAATCTTACGCA 7014  
 Db 65906 GTACTCAGTTTATAGATCTCCGACGAGGACGCTGACGAGTTCTTGGTGAATCTTACGCA 65965  
 QY 7015 GCCCGATCCACAGGGGAGTACTGAGTTAAAGTCCCGAAATCTATCGTTGGTAAT 7074  
 Db 65966 GCCCGATCCACAGGGGAGTACTGAGTTAAAGTCCCGAAATCTATCGTTGGTAAT 66025  
 QY 7075 TTTGTGAGTCTCTTTAGACAGTAACCAATGTTGCTCCGACACTGCGGTACGATGACCG 7134  
 Db 66026 TTTGTGAGTCTCTTTAGACAGTAACCAATGTTGCTCCGACACTGCGGTACGATGACCG 66085  
 QY 7135 CTCTCTTGTCTGTGTCATGAGCACCTGGAGTATCTTAGACAGGTATAGGCGCGTTC 7194  
 Db 66086 CTCTCTTGTCTGTGTCATGAGCACCTGGAGTATCTTAGACAGGTATAGGCGCGTTC 66145  
 QY 7195 CACTCTAAGGGTTCCCGCTTGTGAGCGGATTTGCGGATTTGACCGTCACTGTTGTCAT 7254  
 Db 66146 CACTCTAAGGGTTCCCGCTTGTGAGCGGATTTGCGGATTTGACCGTCACTGTTGTCAT 66205  
 QY 7255 TTCCAGGACTATCATCCGAGTGGGATGAGTGGAACTTGTACATATGATTTAAAGATG 7314  
 Db 66206 TTCCAGGACTATCATCCGAGTGGGATGAGTGGAACTTGTACATATGATTTAAAGATG 66265  
 QY 7315 GTCTTTCCCGTAATGCGACTCATGTGCTGTTTAACTGTTTCAAGGGTGTAAAAAAT 7374  
 Db 66266 GTCTTTCCCGTAATGCGACTCATGTGCTGTTTAACTGTTTCAAGGGTGTAAAAAAT 66325  
 QY 7375 CAGAGTCTCGGTGAGCAACTTTTTCGACTCTCCGACCGCTCAGGAGTAAGCTGTA 7434  
 Db 66326 CAGAGTCTCGGTGAGCAACTTTTTCGACTCTCCGACCGCTCAGGAGTAAGCTGTA 66385  
 QY 7435 CCAGCGGTGCTTCTTATAACGAAGCATACGTCAAAACATGACCCACGTGACCTAAA 7494  
 Db 66386 CCAGCGGTGCTTCTTATAACGAAGCATACGTCAAAACATGACCCACGTGACCTAAA 66445  
 QY 7495 TAAGTCTACAGAAAGTCTGTGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCAT 7554  
 Db 66446 TAAGTCTACAGAAAGTCTGTGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCAT 66505  
 QY 7555 CTCGGGCTCATCGCGATCTGAATCTTTTAACTGTTCTGCTCTACCGGGTTCCCG 7614  
 Db 66506 CTCGGGCTCATCGCGATCTGAATCTTTTAACTGTTCTGCTCTACCGGGTTCCCG 66565  
 QY 7615 GGTTAAACCCCGGGCCCGCACTGTAAACGTCCCGTGGCGAGATGCGCCACAGGGCAT 7674  
 Db 66566 GGTTAAACCCCGGGCCCGCACTGTAAACGTCCCGTGGCGAGATGCGCCACAGGGCAT 66625  
 QY 7675 CGTGAGATTTCTACCGAGAGGGCCCTTGTGTACCGAAAGATGTTCTATACGCTGTC 7734  
 Db 66626 CGTGAGATTTCTACCGAGAGGGCCCTTGTGTACCGAAAGATGTTCTATACGCTGTC 66685  
 QY 7735 CTGAGGGGGTCTTTATACCGCCCGAGTTGGCCAGAGTTCTGACAGAGTCTACCGGA 7794  
 Db 66686 CTGAGGGGGTCTTTATACCGCCCGAGTTGGCCAGAGTTCTGACAGAGTCTACCGGA 66745  
 QY 7795 AATGAAGCCAAAGTTTGGGGGCTGGCTCTCATATATGTTGAGGGGGCTATCAT 7854  
 Db 66746 AATGAAGCCAAAGTTTGGGGGCTGGCTCTCATATATGTTGAGGGGGCTATCAT 66805  
 QY 7855 GGCTATTGCCGACTCGGTCTCTGCTCACTTATAATACCTCGATGCCAGGGAATACTGGA 7914

Db 66806 GGCTATTGCCGACTCGGTCTCTGCTCACTTATAATACCTCGATGCCAGGGAATACTGGA 66865  
 QY 7915 ACTTAGACTCAAGACACTGTGCAAACTTGTTTTTCAACCCACTTTTCTTCAAGAAATCTT 7974  
 Db 66866 CTTTAGACTCAAGACACTGTGCAAACTTGTTTTTCAACCCACTTTTCTTCAAGAAATCTT 66925  
 QY 7975 AGCCCTATGACAGTACTGCGCAACCGGAAGATGGTTCCTGACAAATATTTTACCATCAC 8034  
 Db 66926 AGCCCTATGACAGTACTGCGCAACCGGAAGATGGTTCCTGACAAATATTTTACCATCAC 66985  
 QY 8035 CGGTACGGCCGAGAAAGAGGCGCTGTGCTGACTTGAAGTACTAGCGGGATGACGTGTCC 8094  
 Db 66986 CGGTACGGCCGAGAAAGAGGCGCTGTGCTGACTTGAAGTACTAGCGGGATGACGTGTCC 67045  
 QY 8095 GGGAAAGCAGCCTTGTCCCGATCTCTTAATCTGCGCAGTATGCGAAACCGGGGCTTCTCCC 8154  
 Db 67046 GGGAAAGCAGCCTTGTCCCGATCTCTTAATCTGCGCAGTATGCGAAACCGGGGCTTCTCCC 67105  
 QY 8155 GGCAACCCCTGTTGACCTCAGTAATGCTTAGAAAAATCCAGAAATCATCTCTCAGCGGCCCC 8214  
 Db 67106 GGCAACCCCTGTTGACCTCAGTAATGCTTAGAAAAATCCAGAAATCATCTCTCAGCGCCCC 67165  
 QY 8215 ACCCTGAGTCAATTTGTCATCAAAACACGACCCAGTCTTGCTCAGTCACTCAGTAT 8274  
 Db 67166 ACCCTGAGTCAATTTGTCATCAAAACACGACCCAGTCTTGCTCAGTCACTCAGTAT 67225  
 QY 8275 TATTACGCCAACCCAGGGCGTTGTTCCCGCAATGTTTATGACACAGTGGAAAGCGGT 8334  
 Db 67226 TATTACGCCAACCCAGGGCGTTGTTCCCGCAATGTTTATGACACAGTGGAAAGCGGT 67285  
 QY 8335 GTACAGAGCATTTACCAACAGGACACAGAGCGCTTATTTGGCCCGCGACATTAACCGGTT 8394  
 Db 67286 GTACAGAGCATTTACCAACAGGACACAGAGCGCTTATTTGGCCCGCGACATTAACCGGTT 67345  
 QY 8395 GACATCTGCGCCCTTGGCCCGCATATCGCATGTTCCCGATGTTCCCGCAGTCTCCGCA 8454  
 Db 67346 GACATCTGCGCCCTTGGCCCGCATATCGCATGTTCCCGATGTTCCCGCAGTCTCCGCA 67405  
 QY 8455 GGTGGAGGGTCCCGGGCGTCTGATTTGCGGAAACCCAGCATGCGTTCCGACGCGCGTT 8514  
 Db 67406 GGTGGAGGGTCCCGGGCGTCTGATTTGCGGAAACCCAGCATGCGTTCCGACGCGCGTT 67465  
 QY 8515 ACCGCCAATGTCCCGCAAGAGATGGAACCGTAGCACAGTGGGAAACGCTCCCGT 8574  
 Db 67466 ACCGCCAATGTCCCGCAAGAGATGGAACCGTAGCACAGTGGGAAACGCTCCCGT 67525  
 QY 8575 AAAAAACCTCCACATCGGAGCGCGTATACGCTCCACTGTTTAAATATACCAATAATAGA 8634  
 Db 67526 AAAAAACCTCCACATCGGAGCGCGTATACGCTCCACTGTTTAAATATACCAATAATAGA 67585  
 QY 8635 CTTAACTGCTCCCGTCAAGGTCCGCGCAGATCCGCGCATATCGCAACACTCCAGAGTC 8694  
 Db 67586 CTTAACTGCTCCCGTCAAGGTCCGCGCAGATCCGCGCATATCGCAACACTCCAGAGTC 67645  
 QY 8695 CCGCATGCGCGCGCTCTCCCGCTTCCCGCAAAACCGCCGCAACCGTCCCGCTAAGAG 8754  
 Db 67646 CCGCATGCGCGCGCTCTCCCGCTTCCCGCAAAACCGCCGCAACCGTCCCGCTAAGAG 67705  
 QY 8755 AAAGCAGCACCGCAGGACGTTGGCAGACAAAGACTGAAAGGGGACGTTTGGGGGGCGCG 8814  
 Db 67706 AAAGCAGCACCGCAGGACGTTGGCAGACAAAGACTGAAAGGGGACGTTTGGGGGGCGCG 67765  
 QY 8815 AACAGTAAACCACTTTCCCGGACCGTCCGGATGCGGTTGCGAGCAGGGCTTATT 8874  
 Db 67766 AACAGTAAACCACTTTCCCGGACCGTCCGGATGCGGTTGCGAGCAGGGCTTATT 67825  
 QY 8875 CCAATTTAAATCGAAAGCTCCAGGATGTAAACCGCGAAACGATCTGGAACCAAAAAACGAGA 8934  
 Db 67826 CCAATTTAAATCGAAAGCTCCAGGATGTAAACCGCGAAACGATCTGGAACCAAAAAACGAGA 67885  
 QY 8935 CCAATGCTAGGGGTATCTTACAGGACTCTATGACTACAGTCCCGCGGCGCATCGA 8994

|                       |  |   |                             |
|-----------------------|--|---|-----------------------------|
| Db                    | 67886  | CGACATGCTAGCGGCTATCTCTACAGGACCTGTATGAGCTACAGTATCCCTCCCGCCCATCGA | 67945                       |
| Qy                    | 8995   | TTCCCCCTCCAGCAAACTCGGACAAATGAGGAGATATTTCCAGAGGTTAGTCCGCCCATCTAG | 9054                        |
| Db                    | 67946  | TTCCCCCTCCAGCAAACTCGGACAAATGAGGAGATATTTCCAGAGGTTAGTCCGCCCATCTAG | 68005                       |
| Qy                    | 9055   | CGGCGACGAGTACGCTTGA   | 9073                        |
| Db                    | 68006  | CGGCCACGGATCGCTTGA  | 68024                       |
| RESULT 3              |  |   |                             |
| AF241163              |  |   |                             |
| LOCUS                 | AF241163   | 1945 bp   | mRNA linear VZL 23-OCT-2002 |
| DEFINITION            | Cercopithecine herpesvirus 17 immediate-early transcription factor                     |   |                             |
| ACCESSION             | Rta mRNA, complete cds.  |   |                             |
| VERSION               | AF241163   |   |                             |
| KEYWORDS              | AF241163.1 GI:13506778   |   |                             |
| SOURCE                | Cercopithecine herpesvirus 17  |   |                             |
| ORGANISM              | Cercopithecine herpesvirus 17  |   |                             |
| REFERENCE             | Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus. |   |                             |
| AUTHORS               | 1 (bases 1 to 1945)  |   |                             |
| TITLE                 | Lin,S.F., Robinson,D.R., Oh,J., Jung,J.U., Luciw,P.A. and Kung,H.J.                    |   |                             |
| JOURNAL               | Identification of the bZIP and Rta homologues in the genome of                         |   |                             |
| MEDLINE               | rhesus monkey rhadinovirus   |   |                             |
| PUBMED                | Virology 298 (2), 181-188 (2002)   |   |                             |
| REFERENCE             | 2 (bases 1 to 1945)  |   |                             |
| AUTHORS               | Lin,S.F., Robinson,D.R., Oh,J., Kung,H.J. and Luciw,P.A.                               |   |                             |
| TITLE                 | Cloning and characterization of cDNAs encoding Rhesus Rhadinovirus                     |   |                             |
| JOURNAL               | (RRV) ORP50  |   |                             |
| REFERENCE             | 3 (bases 1 to 1945)  |   |                             |
| AUTHORS               | Lin,S.F., Robinson,D.R., Oh,J., Kung,H.J. and Luciw,P.A.                               |   |                             |
| TITLE                 | Direct Submision   |   |                             |
| JOURNAL               | Submitted (02-MAR-2000) Biological Chemistry, School of Medicine,                      |   |                             |
| FEATURES              | UC Davis, 4645 Second Avenue, Rm. 2400, Sacramento, CA 95831, USA                      |   |                             |
| source                | Location/Qualifiers  |   |                             |
|                       | 1..1945  |   |                             |
|                       | /organism="Cercopithecine herpesvirus 17"  |   |                             |
|                       | /mol_type="mRNA"   |   |                             |
|                       | /isolate="17577"   |   |                             |
|                       | /specific_host="Macaca mulatta"  |   |                             |
|                       | /db_xref="taxon:154334"  |   |                             |
|                       | 65..1798   |   |                             |
|                       | /note="ORF50"  |   |                             |
|                       | /codon_start=1   |   |                             |
|                       | /product="immediate-early transcription factor Rrta"                                   |   |                             |
|                       | /protein_id="AAK28330.1"   |   |                             |
|                       | /db_xref="GI:13506779"   |   |                             |
|                       | /translation="MPTDNDGQVKKLRRCNVNFCGLSEPLRDLKLVHVLVLINEAVK                              |   |                             |
|                       | KHPDRLNKLTEVCRCINVECASLGPIGLIADLNFLNLCYIGSRVKTGKATCN                                   |   |                             |
|                       | VPACCAQGIIVRLTALUCCITSEMFISACSGVIPPQLARVLHVDVTAEMKAKCLGA                               |   |                             |
|                       | WRLLICRRPIAIDSVLVYNTNDAEGLDLRLKALCKVQFIQRIILAPMLLI                                     |   |                             |
|                       | ANGAMVDNPYITLTGAERKRPVVTGSGMTCPGSLVPDSLLILFVCEPGLLPAPLV                                |   |                             |
|                       | DLNSNPNPBIILSAPLSQFVITNTHPSLPOSVITITPTQWVPQCQFMDFTWKAQV                                |   |                             |
|                       | SIHQACTPTLAAATGTSAAAPPHIACSPVAGTSQVSGAGVDGCGKPCVCPQPAL                                 |   |                             |
|                       | ESNNAGSPFFAETAATVPAKRQKPRDEADVAKLGDVKGAAATVNHPPFGPSGMVRE                               |   |                             |
|                       | QGFLDLEISTDYTANASGPKNDLMLAAILDLYLGLQFPPIADSPSSNSDNEEIFE                                |   |                             |
|                       | VSPSPSGHGSP"   |   |                             |
| ORIGIN                |  |   |                             |
| Query Match           | 18.9%  | Score 1714.4;   | DB 14; Length 1945;         |
| Best Local Similarity | 99.9%  | Pred. No. 0;  |                             |
| Matches 1715;         | Conservative   | 0; Mismatches   | 1; Indels 0; Gaps 0;        |
| Qy                    | 7358   | CAGGCTGTAAAAAATCAGAAGTCTCTCGTGGACAACTTTTGGGACTCTCCGAACG         | 7417                        |
| Db                    | 83   | CAGGCTGTAAAAAATCAGAAGTCTCTCGTGGACAACTTTTGGGACTCTCCGAACG         | 142                         |

|   |   |
|---|---|
| JOURNAL<br>MEDLINE<br>PUBMED<br>REFERENCE | (human herpesvirus 8) genes in Kaposi sarcoma<br>Proc. Natl. Acad. Sci. U.S.A. 93 (13), 6641-6646 (1996)<br>96270595<br>8692871<br>4 (bases 122794 to 123567; 123809 to 127297; 129372 to 130400;<br>130551 to 134441)  |
| AUTHORS                                   | Ceserman, E., Nador, R.G., Bai, F., Bohenzky, R.A., Russo, J.J.,<br>Moore, P.S., Chang, Y. and Knowles, D.M.  |
| TITLE                                     | Kaposi's sarcoma-associated herpesvirus contains G protein-coupled<br>receptor and cyclin D homologs which are expressed in Kaposi's<br>sarcoma and malignant lymphoma  |
| JOURNAL<br>MEDLINE<br>PUBMED<br>REFERENCE | J. Virol. 70 (11), 8218-8223 (1996)<br>97048116<br>882957<br>5 (bases 1 to 137508)  |
| AUTHORS                                   | Moore, P.S., Boshoff, C., Weiss, R.A. and Chang, Y.   |
| TITLE                                     | Molecular mimicry of human cytokine and cytokine response pathway<br>genes by KSHV  |
| JOURNAL<br>MEDLINE<br>PUBMED<br>REFERENCE | Science 274 (5293), 1739-1744 (1996)<br>97094384<br>8939871<br>6 (bases 1 to 137508)  |
| AUTHORS                                   | Russo, J.J., Bohenzky, R.A., Chien, M.-C., Chen, J., Yan, M.,<br>Maddalena, D., Parry, J.P., Peruzzi, D., Edelman, I.S., Chang, Y. and<br>Moore, P.S.   |
| TITLE                                     | Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus<br>(HHV8)  |
| JOURNAL<br>MEDLINE<br>PUBMED<br>REFERENCE | Proc. Natl. Acad. Sci. U.S.A. 93 (25), 14862-14867 (1996)<br>97121480<br>8962146<br>7 (bases 1 to 137508)   |
| AUTHORS                                   | Russo, J.J., Bohenzky, R.A., Chien, M.-C., Chen, J., Yan, M.,<br>Maddalena, D., Parry, J.P., Peruzzi, D., Edelman, I.S., Chang, Y. and<br>Moore, P.S.   |
| TITLE                                     | Direct Submission   |
| JOURNAL                                   | Submitted (17-OCT-1996) Dept of Pathology/Genome Center/Division of<br>Epidemiology, Columbia University, 650 West 168th Street, New York,<br>NY 10032, USA   |
| REFERENCE                                 | 8 (bases 1 to 137508)   |
| AUTHORS                                   | Russo, J.J., Bohenzky, R.A., Chien, M.-C., Chen, J., Yan, M.,<br>Maddalena, D., Parry, J.P., Peruzzi, D., Edelman, I.S., Chang, Y. and<br>Moore, P.S.   |
| TITLE                                     | Direct Submission   |
| JOURNAL                                   | Submitted (02-MAY-1997) Dept of Pathology/Genome Center/Division of<br>Epidemiology, Columbia University, 650 West 168th Street, New York,<br>NY 10032, USA   |
| REMARK                                    | Sequence update by submitter  |
| COMMENT                                   | On May 3, 1997 this sequence version replaced gi:1718251.   |
| FEATURES                                  | Location/Qualifiers   |
| source                                    | 1..137508<br>/organism="Human herpesvirus 8"<br>/mol_type="genomic DNA"<br>/db_xref="taxon:37296"<br>/cell_line="BC-1; virally infected primary effusion<br>lymphoma derived cell line"<br>/notes="sequence derived from overlapping lambda phage and<br>cosmid clones isolated from libraries; A prototype<br>sequence for the terminal repeat (TR) unit is deposited<br>under GenBank Accession Number U75639. The left junction<br>variant TR sequence unit immediately preceding base pair 1<br>of the LUR is deposited under GenBank Accession Number<br>U75700. The remaining LUR sequence extending from this<br>sequence to the right TR unit is not available due to<br>difficulties in cloning and sequencing of this region" |
| misc_feature                              | 1..137508<br>/notes="long unique region LUR"  |
| CDS                                       | 105..974<br>/notes="ORF K1"<br>/codon_start=1<br>/protein_id="AAC57081.1"<br>/db_xref="GI:1718252"<br>/translation="MFVLYVCSLAVCFRGLSLSLSSPNLCQGVSTPTVLTCLSNAL<br>SUFISWICNNTRLLRTERVILDTACNFTCVQSGHRQS1WITWPAQFVLQTLCAQ  |

CDS

PSNTVCGQHVTVLYCSTSGNNVTVVHLPNGRNETVSQTKYVNFILMSQTEGCVTCSNLS  
LSSLSNRIICFWARCANITPETHTVSVSTSTPTSTNSLVKIIHATTHRDVVVKEA  
LSTHFIHEVHPLFVMTLVALIGTMCIGILGTIIFAHCQKQRDSNKTVPQQLQDYLYSLD  
LCTDEYCPVDVWY"  
1142. 2794  
/note="ORF 4: The HVS ORF 4 homolog has alternatively  
spliced messages encoding membrane bound and secreted  
proteins; complement binding protein vCBP homolog"  
/codon\_start=1  
/protein\_id="AAC57082.1"  
/db\_xref="GI:1718253"  
/translation="MAFLRQTLWILFTWMTVIGDNEKCSOKTLIGVRLKMSRDGDA  
VGETVELRCRGVTVYARNTATCLOGWSEPTATCNKSCPNPQKNGKHPKIKN  
QDALKYGNISVYCNIEGYVLGVEYRYCMIGASQMAWSSPPCEKCKHRPKIKN  
GDQFKYDNDYDAHFCNEGIVLGHPSIACVANNVTSNMPTELACAGCKPSPVT  
HGYQFSESTYKRGKSVTACNDGVEFLSPKSTETTCNTEWDPCLPCLLVEDIDPNN  
SNPGLKPTNEKNGVNFORSYPTPEPTEDTATACTNCEQPFPKILPTDSGQ  
DATTSTNITKQEDKTIENQPNTHIISALTSMKAGNFTNNNTSLDHTASTPTSD  
EATTEIPSQVPTNYNAPRTLTIHSEGSNKTSEKATSTLSHNSHNGKDTGQI  
VTLNKTQTLSPKTPNSQAKSSTKPRVETHNKTSTNPAISLSDSDADVQRPREPTL  
PIPRFPASKNRYLEKQVLIGLTLTAVALTGCLTILFHYLPFR"  
3210. .6611  
/note="ORF 6: ss DNA binding protein ssDBP homolog; EBV  
BALF2 homolog"  
/codon\_start=1  
/protein\_id="AAC57083.1"  
/db\_xref="GI:1718254"  
/translation="WALKGQQTLEENIGSAATPGCGYLYAYLTNFPPIGEASLLNG  
YPEAKVFLPILHGLTVESDFPLNVAKVAHKIDATIASVKLTSVTRSAIYFHNTHLPQ  
PIFOGKGLKICRESLFGFTVQHQKGLTWSPEACPQLPCANFIAVAVITBGF  
KERLYGKLVPSVQPTVFIHGQAFKIPKIDLEDLFGFSQAELCRFPYVITSRYLH  
DSIFGTIAQALRVKDVSTVQAASRGVFMHQYKIPKIQAKDFPQCAKSTDCGSLTW  
IDSLVAYLGMVSGYLSFTGEGPDCEVLYNTDWPIKENCETPDARLAEVWMAEQALH  
TGAQLPAANSVLYLTRVAKLPQKNRGDANWYNSFVLOHGLGYLSEATVWENGASAFK  
GVPSMDGSSYTIQHLAYASFSFPHLLARMCYIQLPHKNTNSGSYNVVDYGTAA  
APSQMDCQCCQCAVCINTFLYRMKDPPFVLSNKRDPVITGTAGTNDNLELGN  
PATFEREEENGLVEDAPKTYVQWLQONTTEKLSMGISEGDALRTLIVDIPSVFKY  
FKGIDSTVEAEALLFCINCMYKLNINFERENIKSVHILQFACNVYQACPVFVLYLYK  
SLTVIQCILTSQMTVEONPAVGVPSWEMLKMFQTMNFKGACFDKAGTIGCEL  
KIVHOSMCDLFTDAAIGWPEAPFELNAHKLTPSKTTSALYLMHKIQTQKNPILPQ  
TOAGKWPASORDSVTVGGPKMFLNAHKLTPSKTTSALYLMHKIQTQKNPILPQ  
VSGEHLTCLNIVVASSQAEEINLVDLPDILTYSAKILNSSLILACQQTQQTAYIT  
LCSLSPVQLVPAEYHPVLGSLSPDEYRAKVAGRSVTIVOSTLQKAVTNGNRLR  
PIITVPLVNYKTSNGNTNFFVNCALNGYFGSRVDNRNLPESVFFKKNVNSMLKRR  
HYMTPLVDLVKRTIGVNGSGFEABAKVRSQNVLEDNDPNEPFTVLELWPKVPRW  
SSCALBEDVYGVYAVLGEVLSLSTVGQAGVPTWTAAGVASVIQDIIDDCLELQ  
FVGPSEPCILQGSQVVEELFPSGPVSLTVGKKRKIASLLSDLDL"  
6628. .8715  
/note="ORF 7: transport protein homolog; EBV BALF3  
homolog"  
/codon\_start=1  
/protein\_id="AAC57084.1"  
/db\_xref="GI:1718255"  
/translation="WAKELAADVAYSALAMDCLLSVADPATLDTKSLTKTLTKGFS  
LHGTLLPLLRQNAHECSGISLELHFWKTMWMPRECALAENCLQKSIIPSCSLWT  
QATNRSVRNFYGNWALSELKLSLINDVEIFFKLSSVFFYCISSGALGELGVLF  
VGKRGISVPFGLPVLNPLFUECLQVCLTNQGLSDPDISSLMVAGDHHIIEGTVRSVL  
PYRGFENELKQLGQTPESIPTPQCSVRQDDEIRSSLMVAGDHHIIEGTVRSVL  
ETNLIVYSSGSHDTCDBRGDSCHSLAFLTHEADMKRRVDVLACLGEGRTPKRPHF  
CFRPLSLTEFCGLFYSQVDTTESLQKCSAFYQOVNYTALQKAFQVFKVLSKLL  
RAGQNLKCTSPSCQSEARQLVGGKPEEVLROAKHREQLYLVQKVADEFKKLSDC  
IRHQHILSQILGRLVGSIYNEASALONHFLHRAQFISLPMODLTVDCTPEPMSK  
YIKNSLYCRLGRBHEVILLTEFKLITQGLSKRHTLFPSPPNVTLACQTEAAGMLPH  
QYKQSEMIVSIFBPMQIBPNFYQSFENQDINHQKRAWEYIRELVLSYLNKPT  
WERELKILLTPQSGSPGPEEPKPAGLTGLYLTFTSAPLVLVDDKYGWIYKQNLVALLY  
HHQLSNHDSQV"  
8699. .11236  
/note="ORF 8: glycoprotein B gB homolog; EBV BALF4  
homolog"  
/codon\_start=1  
/protein\_id="AAC57085.1"  
/db\_xref="GI:1718256"  
/translation="WTPRSRLATGTVILLVCFCAGAHSRGDTFTQSSSPPTPGSSS

CDS

QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTTACCGCCAGGGGGCGGTTCCCGTGCAATAGTATGC 720  
 Db 65552 TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTGGCTTGCATAGTCTGC 65611  
 QY 721 GTCCGTTCCGCAACAGCAAAACGCGTTCCAGTCAACTTAACAACAGCGGAGCAAAAG 780  
 Db 65612 GTGGGCTCTCCACCCAGACGACGCGCTTTCACTGGTCTTCAACACACAGCAGCAGAGA 65671  
 QY 781 ACGGAGATATCGTCGCGAGAACATCTTAACATTCATGATAGGGAAGAAGTCTATCT 840  
 Db 65672 AACGAGATATCTGCTGTGATATATGCTCACTTCTTATGGGAAACGTTGAGTTGA 65731  
 QY 841 GAGTACGTACATTTGGAGAGAACTGGGCGCTGTTTATATAACAACAAGCGCTGCAACGAT 900  
 Db 65732 GAATATATTAGGCTGACAGAAATTGGGCGCTATTTATATAACAATAAGCGCTGTACGAT 65791  
 QY 901 CTCAGTTCGGACACCTTCTAATAATTTAGAGTAATACTTCCCAATCTCTGACGAATC 960  
 Db 65792 CCCAGTTGGTCACTTGTCTGAAGACCTTAGAATAAATCTAGACATATCACAGAGTTA 65851  
 QY 961 ATGAGTTACGTACAGAGATTTGTCTTCTTAAAGTAAGATATATGGATCTTTTGAATATC 1020  
 Db 65852 ATGACTATATAGATAGTTTGTGTTCCGAAGATAGATTTCTGGACCCGCTCGAGTAT 65911  
 QY 1021 ATTGGCTGGACCGCATCTTTTGTTCATAGCGAGTAAAGCGTATCTTAACAACCTA 1080  
 Db 65912 GCAGGGTGGACAAGACTTTCTATCTCACACGAGGAGTGAAGTCTTTTCTGGCAACGCTG 65971  
 QY 1081 CACACATGTCTAACGCTAGGGGCGGATACAGGGACACGAAGCTCTTTTACCTGTCCCGTG 1140  
 Db 65972 CACACCTGCTGCGAGTAATAGGATGCTGTCTCCACAAGCTTTTCACTCTGCCAGTG 66031  
 QY 1141 GTGTGGAGGTGTTTGTGAAGCCGTTTGAAGAAATAAAGCGGCGCTCAACCTCACCAC 1200  
 Db 66032 GTCTGTAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGGGCGGTAGGCTCTCACAC 66091  
 QY 1201 CTCACGCTGACGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTATTCGAGTTT 1260  
 Db 66092 ATGACTCCCATAGAAATGGGTAAACAAAATCTTTTCAAGCTAAGTAACTCTCGAGTTT 66151  
 QY 1261 GTGGACAGGACATGTCATPAGTGGCCAGGAATCCACTGAAGCTTTCGACCCAGGTTACC 1320  
 Db 66152 GCTGATCAGGACATGGCTGTGGTGGACCTATATCACAGACGCGTCCACAGATCAC 66211  
 QY 1321 TTTATCACCAAGTTGTTTAAACACGCGCTGCTTTTAAATGGAATAACAAAATGTC 1380  
 Db 66212 TTGCCACTAAATTTGTCAAAAACAGCTATGCTACCCCTTACTGGAAGAACCAAAAATGT 66271  
 QY 1381 ATATGTGGGTTTCAAGGTACGTACTTTTGAATTCAAAAGAAATCCTAGACAGCGAATCTTC 1440  
 Db 66272 ATATGCGGTTTCAAGGTCATACCAAGATTCAAGTCCATCTAGACGGGAGCTATTT 66331  
 QY 1441 GTGGAGACGATTCGCAAGCCGTCGGAATACGTATACGGTTTCTTAACACGCTACTG 1500  
 Db 66332 ATCGAAAGTCAATTCGCAAGATAACCCCGCTTATGTGTACAGTTTCTTATGACCTGCTA 66391  
 QY 1501 TACAACGCCATGTACTCGTTTACGCGTACGGGTGACTAGGTCAACAGAGAAATATCTG 1560  
 Db 66392 TATATGCCATGTACTCATTTTACGCGCATCGGGTGAAGCAGGGCATGAAGATTTCTC 66451  
 QY 1561 CAAGACTAAATTTGCAACCCCTCCCGCGCTCTGCGCAACCGGGGTGTAGACCTTCAA 1620  
 Db 66452 AGGACCTCAGGAACTCGCGGTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 66511  
 QY 1621 ACGGTTGCTGAAGATTTAAACCTGGAAGACGACATCTTTACACAGCTGTGCACTGCCCG 1680  
 Db 66512 GTTCTGGGCGAGGAGGGGACACAGATG-----CCTTCTACTCACCGCAGCTCCCA 66565  
 QY 1681 CCACCCGCGGATACCTCTCCCTCCAGTTTGTGTGACACGTAAGTCTGCGCCCTTAAAGGAC 1740  
 Db 66566 CCATCCCCCAACCCAGCGGCTCTTCCCAACACTGTGTGCGCTTATTAATCTCCGGGGCAAGAA 66625  
 QY 1741 GTGTTGCGCTCCAGAAATAAGGTGGCGTGTGCTGTGTTGCGGGAGTTTGAGAAGGAA 1800

Db 66626 CTAATCTGCAACAGCTGGCCCTGGCACGCGACACTTTGGTGACGAGTTCTCTCCACTCC 66685  
 QY 1801 ACGTTTTCGCGTTTACGGTTTAACTGGTGAAGGACGAGTTGACTTTGTCTCCCT 1860  
 Db 66886 GATTTTTCACAGTTTACGGTGAACATCGTGTGGAGATCGGTGAGACTTTGTCTCACT 66745  
 QY 1861 TCAGAACGTCTCAACAGCGCTGTTGGCGTTTGCATCGACCGTTGAATCGTATAAATAAAG 1920  
 Db 66746 TCCCGCGGCTCCACGCTCTAGTGGCATAGCATCCACTATAGACACTATATAATCCAG 66805  
 QY 1921 GGGTACACGTTTACCGGTAGCGTTCGTCGCTGTGACGGTTCAGGCTCAGTGATGAC 1980  
 Db 66806 GGAATATAGTTTCTCCAGTGAAGATTCGCGCTCCAGGAGACAGCGCTCAGCGAGGAC 66865  
 QY 1981 CTCAGGAAGATGCGCTCCCTCGTGTGAGACTCTAGCGGTTTATCGCGTGCCTA 2040  
 Db 66866 CTGCGCAAGATGCGCTCCCTCATAGTTGTCCAGACTCATCGGGTTCAITTCGCTGCTG 66925  
 QY 2041 GAGATTAACATAACCAATTTGACCGAAACCATGGAGACGGGAGCTTTTCCAGTGTGC 2100  
 Db 66926 GAAATAACGTCACCAAGATGACAGACCCCTCGAAGTGGCGAGCTGTTTAAACATATGT 66985  
 QY 2101 TGTCCGGGGAATATGGGTGAGCTCAAAATTTAGCCATGACCATCTGTAAGGACACAGGA 2160  
 Db 66986 TGTCCAGGGGACTACGCTATCAGTTCTAATCGCTATGACCATAGTGAAGGACACAGGG 67045  
 QY 2161 ATGTGTTGGAGCGAGTAGCGGTAGTATTTGGTTCACCAAGACGTCAGACAGCCAC 2220  
 Db 67046 GTTTCACTAAGTGGGTGGCCATATCGTTCGGCAACCAACCGCAATATCAGAGCCAGTCTA 67105  
 QY 2221 GTGTATGTAGCAATATCAAGGGCTGTTAACTCAAAATTTATTTGGTTCATGGACAGCAACCCC 2280  
 Db 67106 GTGTATGTGGTGTATTCAGGGCCATCGACGCTGTTTACCTGTTAATGGACAGTAATCCC 67165  
 QY 2281 CTTAAACCCCTCTCAGAGAACAGTGTGATTAACACTCCGCAAGCATATATAGTCCGCGCC 2340  
 Db 67166 CTTAAGCTAATGAGCCGCGGTGACGCCAGTCCCATCTCTCAAAGTACATCATCAAGGCC 67225  
 QY 2341 CTCACACACCAACACAAACCTCATCTACTAA-----ATAAATAAACAAGCAAAATGGAATAATTTGTG 2373  
 Db 67226 CTATGCAACCCCAAGACTACTCTGATCTACTGACCCGTCACCCCTCTCTTAGGACACTGAT 67285  
 QY 2374 -----ATAAATAAACAAGCAAAATGGAATAATTTGTG 2406  
 Db 67286 GTGTTTGGGAATAAAGCATGAGACTTGACACCTATAATGGTCTGTATTTGACACCATTTCT 67345  
 QY 2407 TTTTATTTCACTCAACACGCGCCACGGATAGTTGTCAATTTCCACACACACGCGGGGGTGG 2466  
 Db 67346 TTAATTTATCACTCCAGCCACGCGCATATATGACCGCTTTCCACACAGGGGTGGCGTGG 67405  
 QY 2467 CCGCCATAGTTTGACGACCAAGAAACGCGCGGCTCTGGGTTAGAAAGTAGACCC----- 2521  
 Db 67406 AGGCCAGGATCGGGTTGGGTGGCTGACCTGGACCCCGCGGTAGTTGTGCTTCTGATG 67465  
 QY 2522 -----GCCCGCGTGTGAGGTCTACGTCTCTTAGGCGCTGGGT 2564  
 Db 67466 AAATCGAGTGGCGGAAGTACTGGGAGATTGGGTGGAGGTGACCTTTGTGCTGACG 67525  
 QY 2565 GTACCAACATAAATCAAGTGTGATCAACGTTGAGTCCCTGCGTGAAGCGCGCTCTTG 2624  
 Db 67526 GAGACAGATCAAGCTACGCGGACGAGGAGTCTCTGCTGTGTCTACTCCCGAGGATA 67585  
 QY 2625 CAGCGGACCAACACCGCGCTGTGAACGTGATCTG-----CAT 2664  
 Db 67586 TAAATATCACGAGACCCACTGCTTTTGGCGCTTAAAGTTGGTGTCTGTGGAGCGCAACA 67645  
 QY 2665 CGTGTGCTGTCTGAAGATGAACACAGGTCAATCAGTGTGCGCGCTGCGCTCAT 2724  
 Db 67646 CATCTCGCTACAGAGAGGCGGTAGACTGCTTTTGGCTTTGCGGCCACGCTCCATGA 67705  
 QY 2725 CCGCGTCTGATTCGPTCCCATGTCTCAACAGAGCCTCGGTGCTGCTGCTGCTGCTG 2784









|       |                    |                             |                                  |       |
|-------|--------------------|-----------------------------|----------------------------------|-------|
| 72081 | ACAAATGGGAGAGCACA  | CTCTGTGCAGTCCGGCGTCTGGGGTTC | TGATTCACAGGGC                    | 72144 |
| 6944  | TGGCGTGGGGTACTCC   | AGTTTTAGGATCTCCG            | CAGACGCGTACGAGTTCTTC             | 7003  |
| 72141 | GCCGTGTGGGGTATTG   | GAGAGCTCAAAACTCTGG          | CAGTCCTTAA                       | 72200 |
| 7004  | ATCTCTACGACCGCGAT  | CCACACAGGGCAGTACTG          | AGTTAATAAGTCCCGCAATCTA           | 7063  |
| 72201 | AAACCTATGCACCCAGG  | TCCTACTAGTGCGCAGCAT         | CGCGTTAATAACACCCCTTATCTTG        | 72260 |
| 7064  | TCGTTTCGTAATTTTCT  | TGGAGCTCCCTTTAGACAG         | TAACCATGTTCTCCGACACTGGGGT        | 7123  |
| 72261 | TCGTTGCCAAGTTTCT   | ACACTGCTCAGGGAATAAG         | CCAAATTCGCCCTAGCCGGGA            | 72320 |
| 7124  | ACGATGAACCGCTCTCT  | TGTTGTCATAGCACCTG           | GAGTATCTTTAGACAGGTAT             | 7183  |
| 72321 | ACCAGTACGGCTCGCTT  | TGTCGTCGTGGACCAATAT         | CTGAATGGTCTTTTGC                 | 72380 |
| 7184  | AGGGCCCGTTCCACTCT  | TAAGGGTTCCGCTTGTAG          | ACGGATTCGGGGCATTCACGGTC          | 7243  |
| 72381 | AGGGTCTCTCAACGTTT  | AGACGGGTAGTGGCACTCT         | GATTCGATTCAGGGTGGCAGCGAC         | 72440 |
| 7244  | AGTGTGTTCAATTTCC   | ACCACTATCATGCCGATG          | CGGATAGTGGAACTTTGCATATGA         | 7303  |
| 72441 | AGGGTATCTAACTCTCT  | GAAGTATCTGATCCAGG           | ACGGTAATGATCCTAAACAGATGG         | 72500 |
| 7304  | TTTAAAGATGGTCTTTT  | CCCGTAATGGCGCACTCA          | TGGTCGTGTTTTA-----AC             | 7352  |
| 72501 | TTGAACAGGTGATCTTT  | AAGGGCCCTTCTCGATG           | TCATTTGTAACACTATGACACGCCAC       | 72560 |
| 7353  | GTTCACAGGGTGTAAAA  | AACTCAGAAAGTCTCGT           | GTGACAACTTTTGC                   | 7412  |
| 72561 | TCTCTCCTTAGGGTAAG  | AAGCTTCGCGCTCTGTG           | TGGAAGCTTCGTGCGCTCTCGG           | 72620 |
| 7413  | AACCGCTCAGGATTAAG  | CTGTACACGGGTGCTTCT          | TATATAACGACATACGTCAAAA           | 7472  |
| 72621 | ACGAATGAAGGCCCACT  | CTACCAAGTGTGTCTCTT          | TATAAATGACCATACGAAACAA           | 72680 |
| 7473  | AACATGACCCACGTCAC  | CTAAATAAGCTCACAGA           | AGAGTCTGTGATTTGTAATGG            | 7532  |
| 72681 | TCTACGATCCAGTACCT  | TAATAGATGGTGGAGAT           | GTGTGCAATTCGGATATGAAG            | 72740 |
| 7533  | AAATGCTCTCTGGGTCC  | CAATCCGGCTCATCGG            | ATCTGATCTCTTTAACCTGT             | 7592  |
| 72741 | AAATGTTCCAACTTGGT  | CGGCTATGTGTCTGTTT           | TACAGACATTAACATGTTTAACTCTTT      | 72800 |
| 7593  | TCGTCTCTACCGGGTTCC | GGGTAAAAACCGGGGCG           | CGCCACCTGTAACTCCCGT              | 7652  |
| 72801 | TCTGTCTTTTTCGTGCCT | CTCGAATAGGACCAAGG           | CGCGCTTACACCTGTGCCAT             | 72860 |
| 7653  | GGCGAGAATGCCCCAGG  | SCATCGTGAGAAATCTC           | ACCGAGAGGCCCTTTGTTGTACCG         | 7712  |
| 72861 | GGCGAGAGCATCCCAAG  | AGCATTTATTCGATCCTC          | CACGAGAGGATCTTATTCG              | 72920 |
| 7713  | AAAAAGATGTTCTATAG  | GGTCTGCTGAGCGGGT            | CGTTATACCGCCCACTGTGGCCAGAG       | 7772  |
| 72921 | AAAAGGCATTTCTGAC   | CGCGATGACGGGGGTG            | AGCTCGCTCCAGCCATATGTAAAGC        | 72980 |
| 7773  | TTCTGCACAGCTCTACG  | CGGAAATGAAGCCCAAG           | TGTTGGGGCGTGGCGTCTCTCA           | 7832  |
| 72981 | TACTACAGAAATATAC   | ATGAAATGAAGGCCAAAT          | GCTGGGGCCTTGAGGGCACTCG           | 73040 |
| 7833  | TATGTTGACGGCGCTAT  | CACTGCGCTATGCGCAT           | TTGCGCACTCGGTCTCTGCTACTTATATACCC | 7892  |
| 73041 | TCTGCAATCGAGGCCCA  | TATATGATTTAACTCTT           | CCCTACTGAAGCTCTACACACGT          | 73100 |
| 7893  | TGGATGCCGAGGAAACT  | TGGAACTTAGACTCAA            | AGCACTGTGCAAACTTGTTTTTCAAC       | 7952  |
| 73101 | ACGATACCGCGGGTGT   | CTCTGAGCAGTCAAG             | GGCCCTCTGCTCTTTGTTTCCAAC         | 73160 |
| 7953  | CCATCTTTCTTCAAGAA  | TCTTTAGCCCTTATCG            | ACTCTGCGCAACCGGAGATGGTTC         | 8012  |
| 73161 | CGGTCTACCTTCGAGG   | ATTAATGCGCGCTCGAG           | ATCATCAACCAAGGTTCAGCTCGCC        | 73220 |

|            |  |   |                            |
|------------|--|---|----------------------------|
| Qy         | 8013   | CTGACAAATATTATTTACATCATCCGGTAGCGCCGAGAGAGGCGCCCTGCTGTCATCGAA  | 8072                       |
| Db         | 73221  | CTGAAACATTTTACAGCATCATCCGGTCTGCTGAGAAACCGCGGCAATATACCATCGGCA  | 73280                      |
| Qy         | 8073   | GTACTAGCGGATGACGTGTGCGGGAAGCAGCGCTGTGCCCGATTCTTAATCTCTGCCAG   | 8132                       |
| Db         | 73281  | AGGTCACTGCACTGTCTATTCAGGAGCGGTCTCATGCCAGATCTTAAATTTTGCCA      | 73340                      |
| Qy         | 8133   | TATGCGAAACGGGGCTTCTCCCGGCACCCCTCGGTGACCTCAGTAATGTCCTTAGAAATC  | 8192                       |
| Db         | 73341  | TCCTGGAGCCAGGACTGTGTGCGCGTCTCCATGGTAGACCTCAGCAGTGTGCTGGCAAAAC | 73400                      |
| Qy         | 8193   | CAGAAATCATCTCTCAGCGCCGCCACCCCTCAGTCAATTTGTCATCAAAACAGCGACCCCA | 8252                       |
| Db         | 73401  | CCCGGTTATTCGAGCGGCCCTCCCTGAGCGAGTTGTGTCATAGCAACCCCATCCCA      | 73460                      |
| Qy         | 8253   | GTCTGCCTCAGTCAGTCAGCATTTATPACCCAACCCAGCGGCTGTGTCGCCGCCAATGTT  | 8312                       |
| Db         | 73461  | ACATGCCGACACCGGTACGATCATCCCTTTAAAC--CATCGGTACAGACCCCGGCT      | 73517                      |
| Qy         | 8313   | TTATGGACACGTGGAAGCGGTGTACAGAGCATTCACACACGACACGAC              | 8364                       |
| Db         | 73518  | TTATTAGTACGTGGAGCGCGGTACAGATATGTTGTAACATCCAC                  | 73569                      |
| RESULT 5   |  |   |                            |
| LOCUS      | U93872   | 133661 bp   | DNA linear VRL 09-JUL-2001 |
| DEFINITION | Kaposi's sarcoma-associated herpesvirus glycoprotein M, DNA replication protein, glycoprotein, DNA replication protein, FLICE inhibitory protein and v-cyclin genes, complete cds, and tegument protein gene, partial cds. |   |                            |
| ACCESSION  | U93872   |   |                            |
| VERSION    | U93872   |   |                            |
| KEYWORDS   | U93872.2   | GI:14627174   |                            |
| SOURCE     | Human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus - Human herpesvirus 8)  |   |                            |
| ORGANISM   | Human herpesvirus 8  |   |                            |
| REFERENCE  | Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.   |   |                            |
| AUTHORS    | 1 (bases 29032 to 30108; 117733 to 118431)   |   |                            |
| TITLE      | Zhong, W., Wang, H., Herndler, B. and Ganem, D.  |   |                            |
| JOURNAL    | Restricted expression of Kaposi sarcoma-associated herpesvirus (human herpesvirus 8) genes in Kaposi sarcoma   |   |                            |
| MEDLINE    | Proc. Natl. Acad. Sci. U.S.A. 93 (13), 6641-6646 (1996)  |   |                            |
| PUBMED     | 96270595   |   |                            |
| REFERENCE  | 8692871  |   |                            |
| AUTHORS    | Neipel, F., Albrecht, J.C., Ensser, A., Huang, Y.Q., Li, J.J., Friedman-Kien, A.E. and Fleckenstein, B.  |   |                            |
| TITLE      | Human herpesvirus 8 encodes a homolog of interleukin-6   |   |                            |
| JOURNAL    | J. Virol. 71 (1), 839-842 (1997)   |   |                            |
| MEDLINE    | 97138401   |   |                            |
| PUBMED     | 8985427  |   |                            |
| REFERENCE  | (bases 123309 to 124082)   |   |                            |
| AUTHORS    | Li, M., Lee, H., Yoon, D.W., Albrecht, J.C., Fleckenstein, B., Neipel, F., U.  |   |                            |
| TITLE      | Kaposi's sarcoma-associated herpesvirus encodes a functional cyclin  |   |                            |
| JOURNAL    | J. Virol. 71 (3), 1984-1991 (1997)   |   |                            |
| MEDLINE    | 97184528   |   |                            |
| PUBMED     | 9032330  |   |                            |
| REFERENCE  | (bases 122660 to 123226)   |   |                            |
| AUTHORS    | Thome, M., Schneider, P., Hofmann, K., Fickenscher, H., Meinl, E., Neipel, F., Mattmann, C., Burns, K., Bodmer, J.L., Schroter, M., Scaffidi, C., Kramer, P.H., Peter, M.E. and Tschopp, J.                                |   |                            |
| TITLE      | Viral FLICE-inhibitory proteins (FLiPs) prevent apoptosis induced by death receptors   |   |                            |
| JOURNAL    | Nature 386 (6624), 517-521 (1997)  |   |                            |
| MEDLINE    | 97242415   |   |                            |
| PUBMED     | 9087414  |   |                            |
| REFERENCE  | (bases 1 to 133661)  |   |                            |
| AUTHORS    | Neipel, F., Albrecht, J.C. and Fleckenstein, B.  |   |                            |



|                            |       |  |       |
|----------------------------|-------|--|-------|
| CDS                        |       | LPHLTVYKQARBELPQIHDIRIPYFVDAPGSLRSLASHPEYVKQHLRVAVDLY<br>FKLHVAVANIQCULFQNNISATVAMLNFIDIPVTFPTP"<br>14500..115756<br>/note="ORF 10, herpesvirus saimiri ORF 10 homolog,<br>conserved in other gamma-herpesviruses"<br>/codon_start=1<br>/protein_id="AAB62594.1"<br>/db_xref="GI:2246469"<br>/translation="MQTEATFILGDWEIVTSNCRCTSSLTGCLPLRSGDYTLRIP |       |
| Query Match                |       | 17.7%; Score 1602.8; DB 14; Length 133661;   |       |
| Best Local Similarity      |       | 51.5%; Pred. No. 0;  |       |
| Matches 4478; Conservative |       | 0; Mismatches 3872; Indels 342; Gaps 21;   |       |
| Qy                         | 1     | ATCGAGAGTCCGTCCGATGGACCAACACACGTCGACCAAAATCCGGGGTTCATCTTGAAC   | 60    |
| Db                         | 65191 | ATGGACAGCTCGGAAGGGTGCACTGATATGGAGAGCCCGACCCCGGGTTCATCTCAAC   | 65250 |
| Qy                         | 61    | ATGACGTCGATGCCAAAGTCAGGAGCGTAGTGAACAGATTGACAGAGTTGTCAAATATC  | 65310 |
| Db                         | 65251 | ATGACATCGATGCCAAAGTCAGGAGCGTAGTGAACAGATTGACAGAGTTGTCAAATATC  | 65310 |
| Qy                         | 121   | ACTACAGCCACCGGAATGGTGTGACGACCTGGCCCTCGATCCGGCTGAAGACTCC  | 180   |
| Db                         | 65311 | ACTACATCGCCCGGAGATGGCTGGTATGACCTAGAGTTGATGCCATCGGAAGACGAA  | 65370 |
| Qy                         | 181   | GGCCCGTCTTCGCGTTTACCGTTTATCTAAATACCGGAACTGCTGGTCTGGGAAAAGT   | 240   |
| Db                         | 65371 | GGCCCGTCTTCGCGTTTACCGTTTATCTAAATACCGGAACTGCTGGTCTGGGAAAAGT   | 65430 |
| Qy                         | 241   | ACCAGATATCGGCCCTGTACCAAATTTAACTGCTGATCACGGGCGCGACCAACATA   | 300   |
| Db                         | 65431 | ACCAGGATATCGGCCCTGTACCAAATTTAACTGCTGATCACGGGCGCGTACAGTGTA  | 65490 |
| Qy                         | 301   | GGCGCAGAACTATCGCTCGCTTAAGAGCTCTGTCCCAAGATCTTCAGCGCTTTT   | 360   |
| Db                         | 65491 | GGCGCAGAACTATCTCCAGGGCTTTAAGTCTCTACTGTCCCACTATATACCAAGCTTC   | 65550 |
| Qy                         | 361   | GGCTTTAAGAGCGGACACATCAATATAGCCGTGAGAAAAGCTCATCAGACCGGAGCGTA  | 420   |
| Db                         | 65551 | GGATTCAGAGCAGACACATTAATATCTGCCAGAGAAAGTGCCCAAGGTAATCTAGTCC   | 65610 |
| Qy                         | 421   | TCCATAGAGCAATTCAGCAACAGGAGCTATCGAAGTATGGCGGTATAGTGGACATT   | 480   |
| Db                         | 65611 | TCCATGAGCAATTCAGCAACAGGAGCTGCTAGGTATGGCCAACTGTGCACCGATAT   | 65670 |
| Qy                         | 481   | ATGAAGAGGTTATGGCGAAAACCAATGGCATGTACGGGACTATATCCAAACGCGAAT  | 540   |
| Db                         | 65671 | ATTCGAGAAATTTATGCCAGAGAACAAAGGGCGAGTATAGCTCCCTCTCTCAAGCGCT   | 65730 |
| Qy                         | 541   | TTTGAACCCCTCTCGAGAAATGACGGGACCGTGTATGAGCTTCGAATATTATTGTAATC  | 600   |
| Db                         | 65731 | TTCAAGACTCTCTGCGGTATGGGTGGAGCCAAATTTGTGGACGAGTAACATTATCGTGATA  | 65790 |
| Qy                         | 601   | GACGAGCGCGAAACCTGTCTCTTACATACTCACACACGCTCGTGTCTTTTACTGGTTC   | 660   |
| Db                         | 65791 | GACGAGCTGGAACTCTCTGTCCTCATATTTTTCAGCGCGTGTGTCTTCTTATTGTTT  | 65850 |
| Qy                         | 661   | CTAAACAGCTGGCTTAATAACCCCTTTTACCGCAGGGGGCGGTTCCGTGCATATATGC   | 720   |
| Db                         | 65851 | TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTGGCGCTTGTGCATATGCTGC  | 65910 |
| Qy                         | 721   | GTCCGTTCCGCAACGAGACAAACGGCTTCCAGTCAACTTACAAACACGCGGACGCAAAAG   | 780   |
| Db                         | 65911 | GTGGGGTCTCCACCCAGACGAGCGCTTTCACTGGTCTTCAACACACGCGAGCAGAGA  | 65970 |
| Qy                         | 781   | ACGAGATATCGTCTGCGAAGACATCTTAACATTCATGATAGGGAAGAGGTCGTATCT  | 840   |
| Db                         | 65971 | AACGAGATATCTGCTGTGATAATGTGCTCACCTTCTTATTTGGGAAAACGTCAGGTTGCA   | 66030 |
| Qy                         | 841   | GAGTACGTACATTTGGAGAGAACTGGCGCTGTTTATAACAAACAGCGCTGCACCGAT  | 900   |
| Db                         | 66031 | GATTTATATAGGCTGGACGAAATTTGGGCCCTATTATATAAACAATAAGCGCTGTACGGAT  | 66090 |

|    |       |   |       |
|----|-------|---|-------|
| Qy | 901   | CTCAGGTTCCGACACACCTTTCTAAATAATTTTAGAGTATATCTTCCATTCTCTGACGAAGTC | 960   |
| Db | 66091 | CCCGAGTTTGGTCACTTGTCTGAAGACCTTAGAATAATCTAGACATATCACGAGATTGA     | 66150 |
| Qy | 961   | ATGAGTTTACGTAGACAGATTGTCGTTCCCTAAAGTAAGATTATGGATTCCTTTAGAATAC   | 1020  |
| Db | 66151 | ATGGACTATATAGATAGGTTTGTGGTTCCGAAGAGTAAGATTCTGGACCCGCTCGAGTAT    | 66210 |
| Qy | 1021  | ATTGGCTGGACCCGACCTTTTGTGCATATAGGAGGTAAAGGCTATCTAAACAACCTA       | 1080  |
| Db | 66211 | GCAGGGTGGACAGACTCTTCACTCTCACAGAGGTGAAGTCTTTTCTGGACACGCTG        | 66270 |
| Qy | 1081  | CACACATGTCTTAACGCTAGGGGGCGATACCAGGGACACGAAAGCTCTTTTACCTGTCGCTG  | 1140  |
| Db | 66271 | CACACCTGCCTGTCGAGTAATAGGATGCTGTGTCCACAAGACTTTTTCACCTGCCAGTG     | 66330 |
| Qy | 1141  | GTGTGCGAGGTGTTTGTGAAGCCGTTTGAAGATACAGCGGCGCGTCAACCTCACCAAC      | 1200  |
| Db | 66331 | GTCTGTGAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGGGGCGTAGGCTTCACAC      | 66390 |
| Qy | 1201  | CTCACCGTGAACGAGTGGGTGACAAAAAAGCTCTTTAAAGTTAAGTAATTTATTCGAGTTT   | 1260  |
| Db | 66391 | ATGACTCCCATAGAAATGGGTAAACAATAATCTTTTCAAGGCTAAGTAATCTACTCGAGTTT  | 66450 |
| Qy | 1261  | GTGACACGAGACATGTCCTAGTCCGACCGGAATCCACTGAAACGTTTCGACCCAGTTTACC   | 1320  |
| Db | 66451 | GCTGATCAGGACATGGCTGTGTTGGACCTTATATCACAGCGCTCCACACATCACC         | 66510 |
| Qy | 1321  | TTTATACCAAGTTTGTAAAAACAGCACCTGTCTTTAAATGAAAAACAAAAAATGC         | 1380  |
| Db | 66511 | TTCCGCCATAATTTGTCAAAAACAGCTATGCTACCTTACTGGAAGACCAAAAAATGT       | 66570 |
| Qy | 1381  | ATATGTGGTTTTCAGGGTACGTACTTTTGAAGTTCAAAAAGATTCCTAGACAGGAACTCTTC  | 1440  |
| Db | 66571 | ATATGCGGGTTTTCAGGGTCATACAAAAGATTCAAGTCCATCTCTAGACGGGAGCTATTT    | 66630 |
| Qy | 1441  | GTGAGACGCTATTCGCAAGCCGTCCGGAATACCTATACAGTTCTTCTTAAACACAGCTACTG  | 1500  |
| Db | 66631 | ATCGAAATGCTATTCGACGATTAACCCGCTTATGTGTACAGTTTCTTGTAGTACCTTGTA    | 66690 |
| Qy | 1501  | TACAACGCTATGTTACTGTCTTTCAGCGCTACCGCGTGTACTAGGTCAACAGAAATATCTG   | 1560  |
| Db | 66691 | TATAATGCCATGTACTCATTTTACCGCACGGGTGAAGCAGGGGCATGAAGAATTCTCTC     | 66750 |
| Qy | 1561  | CAAGACTAAATTTTGACCCCTCCCGCCGCTCTGGCAACCGGGCGGTGTAGACCTTCAA      | 1620  |
| Db | 66751 | AGGGACCTCAGGGAACTGCGCGTGTCTCAAGAGCTGATCTCTGAGTGAAGTCTCCAGGAC    | 66810 |
| Qy | 1621  | ACGGTTCTGTAAGAGTTAAACCTGGAAGACGACATCTTTTACCACGTTGTGAGTCCCCCG    | 1680  |
| Db | 66811 | GTTCTGGGGCAGGAGGGGACACAGATG-----CCTTCTACCTCACCGCCAGCTCCCA       | 66854 |
| Qy | 1681  | CCACCCGCGGTATCACTCCCTCAGGTTTGGTTCGACACGTAAGTCTGCGCTTAAAGGAC     | 1740  |
| Db | 66851 | CCATCCCCACCCACGCGGCTCTTCCAACTGTGTGCTTATCTCTCGGGGCCAAGGAA        | 66924 |
| Qy | 1741  | GTGTTTGGCTCCAGAAATAAGGTGGCGTGTGCTGTTTGGCGGGAGTCTTGAAGAAGAA      | 1800  |
| Db | 66925 | CTATTCTGCAACAGCTGGCCCTGGCAGCCGACACTTTGTGTGACGAGTCTCTCCACTCC     | 66984 |
| Qy | 1801  | ACGTTTTCCGGTTTACGGTTAAATGTCGTAAAGGAGCGAGTTGACTTTGTCTCCCT        | 1860  |
| Db | 66985 | GATTTTTCAAGCTTTTACGGTGAACATCGTGGTGGAGATGGCGTGGACTTTGTGTCCACT    | 67044 |
| Qy | 1861  | TCAGAACCTCTCAAGGGCTGTGGCGTTTGCATCGACCGTTGATTCGTATAAATAATTAAG    | 1920  |
| Db | 67045 | TCCCCCGGGTTCCACGGTCTAGTGGCATACGATCCACTATAGACACTATATATCCAG       | 67104 |
| Qy | 1921  | GGGTACAGTTTTTACCGGTAGCGTTCCGTGCTGTCTAGGGTTTCCCACTCAGTGATGAC     | 1980  |
| Db | 67105 | GGATATAGTCTCTCCCACTGAGATTCCGCGCTCCAGGAGGACAGCGCTCAGCGAGGAC      | 67164 |



Db 69322 TACTGAAATTGGGACGGAAGCCCTCTAGCTCTCTAAGAATGCTTCTCAAACTGGGTGGAA 69381  
 Qy 3957 TACTGTAGTCCGAGCAACGCTAAAGCTAGTCCGGTGGCTTACCGCGATGGTAAGGAT 4016  
 Db 69382 CCTGACATTCGGGATCCACACTAAACGCCAGGCCAGTAGCTTGGCCCTTGGTACGGGT 69441  
 Qy 4017 CTTGGCCAAAGATCAAGACTTTAATATCTCCGGTTCCGAAACAAATACGACCCACCAATAA 4076  
 Db 69442 CTTGGCCAAAGATCAAGACTTTAATATCTCCGGTTCCGAAACAAATACGACCCACCAATAA 69501  
 Qy 4077 TTTCTATCTATCGGTGGATTAATCAGCGTTAGCTTCTATATATCCATAACCGCTTCAGCA 4136  
 Db 69502 GCTTGTCTCTGGGGGATACACTGTGTGGTTAGCTTATAGTTCCCGAATCTGTCTGAGCA 69561  
 Qy 4137 GAGCCGAGCTTCTGTTTTAAATAATGGAAAGATTTTAAATAATCAAGCCATGAGTCG 4196  
 Db 69562 GCGAGAGCACTTCTGTCTTTCAGAAATGATGAGAGGCTCAGAAAGAAATCCACTTAGGTG 69621  
 Qy 4197 TAAGTAGCA-----AGTTTGTGTGAGGGTCTTCTAACACT 4234  
 Db 69622 CCAGTAACAGATCCCGGTGCTGACCCCTGACCTGATGATGGGTGCCCCCTAAAGACCG 69681  
 Qy 4235 TCTGGTGACATTTTACTCCAAACTATTTGTTTAAAGCAACCCCTCCATAAATCCCGCGTC 4294  
 Db 69682 TCTGTTGCAACCATCGGTCCATGTTGAATTTATTTTCCCTTTGACCTGGGTGGCTCTC 69741  
 Qy 4295 TGTGTAATTAATTACGTAATAGTGAACCTTTTGTACTCTCTGTCTGAGAAATCCACAGAG 4354  
 Db 69742 CGGCTGCTGCTTTTAGCCGAGTCTGACTTCCGCTAACAGAACTCTCCGGTTTCATGGCC 69801  
 Qy 4355 GTGCGCCGATTAAGCAGTTTCCGCTTCACTACCGTTTATGTCAAGT-----CT 4404  
 Db 69802 TTTCCACGCTTATATATTTATTTTACGTTGTGTAATAGAGCTATCTGCAGTGTGCGG 69861  
 Qy 4405 TCAAACTCGGCTAAAACTCATCAAGACTATATTCATATATTAATCAACAATGAAGCTCT 4464  
 Db 69862 TTAACACCTACAGTATAGGCGCTCAAACTTCGTTGTAATATACCAACCAACCTCAGGTTT 69921  
 Qy 4465 TCTCTGGGACACTGTAATTTAATTTACTGTACAAATCTTTATAAAGCCCTTAAGTGA 4524  
 Db 69922 TCTGCGACGCCAGGACCCCAATCTTCGAACACCGCGACTAAATAATGACCTCAGATTA 69981  
 Qy 4525 AATCCATTTACACATATTTCTGTGTAGAGCGATCAGCTTTATATCGCAATGTTGCTGTA 4584  
 Db 69982 AACCATTCAGCATGTTTCCACGGTAATGTCGCCCTGTTTGTCTCGAGCTTGGCTATA 70041  
 Qy 4585 CACAGCCCTGGACATAAAGCTAAATTTGTTAATGAAGAAGCAATCTCAAGCGCTCC 4644  
 Db 70042 CAGACCCCGTTGCAAGTATTCGATCGCGGAGTGGATAGTGGACCGCAAGAACAA 70101  
 Qy 4645 GGAATATAAAGTCACTAAGTGGTATTTATTTTACACATCGAATTTACTACAAGCCCGTA 4704  
 Db 70102 GGCAGGTAGGCTGCGGATGCTGAAATTCGCAACATGTTAAGGCGAGTATGGTGA 70161  
 Qy 4705 CAACATGTAATTAATAAATAATCCACATGTAATTAATAAATTAAGGTATACAT---ACTT 4761  
 Db 70162 GATGTGACCAATAGGTGGTCCACAGACCGGCAATAGCGCAAAAGATCCCATGGGGCAA 70221  
 Qy 4762 CTCATGTTTATACATAATATGAACGTAATGCTGTACTTATATAATATGATAGAAATTA 4821  
 Db 70222 ATCCGGGTTTACCCCTGTGTGCTGGTTCGGTCTCCCGAGGAGCCCTTCGGTAA 70281  
 Qy 4822 TGCCAAATACAGTCTATGTAAGTATATTTTAAAGGCATTAATCAACATCAATAATATAAC 4881  
 Db 70282 TATCTGTTTTATATAGTGGGTTTCAAGCATGCGGAGTCCCGACTAATGAGGACAATTA 70341  
 Qy 4882 ACAAAATGACGCTTCCGGATATCCGCCAGCTGTTAAACCGAATAATACAGAAATGACTA 4941  
 Db 70342 CTGAAATGACCTTTTCGCGACAGGGGGTGAAGTCTTATTTCCACAGCATATCTTCGGG 70401  
 Qy 4942 CACAAACACACTGAAACCAATTTTATCTCAA-----CATATGCAACAATATT 4990  
 Db 70402 GAAAAATACCACTCTTTAATTTCCGTGGGAGACGATGGGGAATGTGCAATTACC 70461

Qy 4991 TCAGGTAACACCAATGTTAATAAATAACGACAGCATGCAATTTTGTAGCTAAGCTCTCTCAA 5050  
 Db 70462 TGACACGGTTCATCAATCATACTCATCGTGGAGCTGTCAACAGCTGTGGCTGAGATTTTCTA 70521  
 Qy 5051 AGACAAATTTCTCTCTCAATTTGATGACTCATCAAC-----CTCGGTTCCTCGGTAAATGTC 5105  
 Db 70522 AAAAGTCAATCAATGAATCATCGGAATCATCAGCACACTCTAAGAACTACTCCATATGCCG 70581  
 Qy 5106 GGAATAATATAGTCCAGGATA-CCCTGCTCATCGCCACAAATTAATTTCCAGGGCTGTGA 5164  
 Db 70582 GGGTGGCGGGGTCCTCGAGTAGTGACCTGCGCATCGGAGACACAGATGATGGGTTG 70641  
 Qy 5165 AAATATCATCCAGTGAATCAGCTCTAATGTCAGAGGATGTTGTTTCTTAATTTTAGCCCAT 5224  
 Db 70642 AAATGTCCATACGGGCGGTGTGCAAGGGTCAAGTCCCATCCCAACACAGAGACCTT 70701  
 Qy 5225 CCAATATGAGTGGGTGACACTTTTGGGACAAGACCCCGTAAATTTGSCACGCTCCACAC 5284  
 Db 70702 TAGATACCCCTCTCCGGCATGTGCGGTAATCCGGGCAAGCAAGCTGTTCTGGATTCC 70761  
 Qy 5285 AAAACGAAAGATGATTTTGTATGCTTACAGAGGGGACGATATTTATTTTATTAACATCAG 5344  
 Db 70762 AAAAGTGGCCAGCGGTACCCAAATCGCAGGGCGTGTATTTATTTTCCACAGAACCG 70821  
 Qy 5345 GAAACGGACCCGCTCAATGACGCACATATGTCATTAAGATTTGAGTCTGTCACACTGGTTG 5404  
 Db 70822 GTTCTCTAATTTGATCATCAGAGGTATCCAAAACCCGGCTTCCACGTTGATCCGGCTTA 70881  
 Qy 5405 AATTGACGGTGAATAAATAATGATTTGCTGTCAATGCTGTAATCTTTTAAACACACGA 5464  
 Db 70882 CCGACAGTCTTTTCCAGGGTTTCTGTGGGGCGCGGAGCTGACTCAAAAAGGTCACTG 70941  
 Qy 5465 GGGCGGAATAAAACGGGTCCCGCGCTAAGGCGGTGAATCGGCGAGCACCGGCTCATTTG 5524  
 Db 70942 CCTTGCCTCATGGGCGGTGGGTGACAGTCCGCATACTCTTCCAGGACACTGGCCATGC 71001  
 Qy 5525 CCATTTCCAGATTTGTTAAATCTGTTGAAACATACGAGTCTPAAATCAAGTATTAACACG 5584  
 Db 71002 ATGACTCCAAACCGTCTCAGTCCGAGGTAAATGTGCTCTATGAAGATGTGGTAGGCCAGC 71061  
 Qy 5585 ACAAGTTCAACAGAGGCAACGGCGTGGTGTTCACCCGACACATCCCGCTCACAAACC 5644  
 Db 71062 AGACGTTCAAAACAGATGAATCAAGCTAAGCTCCCGCGAATCCACATCCCAACAGG 71121  
 Qy 5645 CGAACTCTTAGAACGCACATCAAGGCTTCAGACAAAATTCGAAACCCGAAACCCCTAG 5704  
 Db 71122 GGATTTGCTCGGTGCTGTATTTAGTCTGGAAATAGAAACTCAGAAAAGACACTGACC 71181  
 Qy 5705 TTCTGAGCACCATCTGCTGTTATAAAGTGGGTTAAGAGAGAGCGGCCACACACTGCTGT 5764  
 Db 71182 CACCAAGGAGAACCTGGCGTCTTGCAAAGTTGATGAGCCCGCAAGAAATGTGTCTCC 71241  
 Qy 5765 TGTGAGGATTCAGCTTAGTGGCTGCGCTGACAGGCTTGTGACAGGCCCGTAGCTGCGGG 5824  
 Db 71242 CGTGGACAAAGAGCTTGGGGGGGAGAGATGGCGCTACAGTGGGTGATTTCTTCTACCA 71301  
 Qy 5825 CGCTCGCGCACTTTGAACAGCGGACGCCAGTTCAGAGCGGAGGAGCGCTTTTTC 5884  
 Db 71302 CGGTCTATCATTTGGTGGCACCCACAGGCTGTTCAGTATCAGCATATAATCTATCTTTCG 71361  
 Qy 5885 AGTCCGCGCATATCTCGGCCATAGAGATTCGCTCGCTGATCTTTTAGGTGCAATATA 5944  
 Db 71362 AGTCAATCCAGATCAAGATCATGTGATGCTGTGCTTGGCATTTTGGCCGCGATGACA 71421  
 Qy 5945 TCTGTTGGGTAAACAAACCTCACTCCCTGTAAAGGGGAATAGGTCCGTTCTTTTATCG 6004  
 Db 71422 TTTCTGTGCCACATATTTTAAACATCTGTAATCTGGAAGTGAATTCAGTCTGTGTTGA 71481  
 Qy 6005 TTTTCTCTGAATCTATAGTTAGATCTTTTAAAGTATTAACCAAGACCTTAAACCCAGAT 6064  
 Db 71482 GCCCCCGGGGAAACCCAGCTGATGCTTCAGGACCCACAGGACCTTAGAACCCCGGT 71541

|       |   |       |
|-------|---|-------|
| 72620 | ACCAGGTACGGCTCGCTTTGTTCGGTGTCTGGACCAATATCTGAAATGGTCTTTTGCAGGTAT | 72679 |
| 7184  | AGGCCCGTTTCCACTTAAGGGTTTCCGCTTTGTGTAGACGGATTTCGGGCAATATGACGGTC  | 7243  |
| 72680 | AGGCTCTTCTCAACGTTTATAGACCGGTAACGTGGCAGTCTGTGATTTAGGGTGGCGACGGAC | 72739 |
| 7244  | AGTGTGTCTATTTACCGACTATATCATCCGAGTCCGGATAGTCGGAACCTGTGATCATATGA  | 7303  |
| 72740 | AGGTAATCTAACTCTCTGAAGTATCTGATCCAGACCGGTAAATGATACCTTAAACAGATGG   | 72799 |
| 7304  | TTTAAAGATGGTCTTTCCCGTAATAGCGCACTCATGTCGTGTGTTTAA-----AC         | 7352  |
| 72800 | TTGAAACAGGTGATCTTTAAAGGGCGCTTCTCGATGTCATTTGTAAATACTATGACACGCCAC | 72859 |
| 7353  | GTTTACAGGGTGTAAATAAACTCAGAAGGTCCTCGCTGGACAACTTTTTCGGGACTCTCCG   | 7412  |
| 72860 | TCTCTCTTAGGGTAAGAAGCTTCGGCGGTCTCTGTGTGGAAGCTTCTGTCCGCCCTCTCGG   | 72919 |
| 7413  | AACCGCTCAGGGAATAAGTGTATACCACCGGTGCTTCTTATAAACGAAGCATACGTCAAAA   | 7472  |
| 72920 | ACGAACCTGAAGGCCAACTCTTACCAGTGTGTCTCTTATAAATGACGATACGAACAA       | 72979 |
| 7473  | AACATGACCCACGTGACCTTAATAAGCTCACAAGAAGTCTGTCTGATGTATTTGAATGG     | 7532  |
| 72980 | TCTACGATCCCAAGTACCTTAATAAGATGTGTGGAAGATGTGTGATTCGGGATTTATGAAG   | 73039 |
| 7533  | AATGTGCTCTCCGTGGTCCCACTCTCCGGGCTCATCGGGATCTGAATCTCTTTAACCTGT    | 7592  |
| 73040 | AATGTTCAAGCTTGGTCCGCTATGTGTCTGTTTACAGACATTAACATGTTTAACTTT       | 73099 |
| 7593  | TCTGCTCTACCGGGTTCCTCGGTTTAAACCCCGGGCGCCGCACTGTGTAACTGTCCTGT     | 7652  |
| 73100 | TCTGCTTTTTCGTGCTCTCGAATGAGGACCAAGCGCGCGCGGTACAACTGTCGCAT        | 73159 |
| 7653  | CGCAGATGCGCCACGGGATCTGTGAATCTTCAACGAGAGGCGCTTTGTTGTACCG         | 7712  |
| 73160 | CGCAGAGGCAATCCCAAGGCATTAATCGGATCTTCAACGAGAGGATCTTATTTCTGCACAG   | 73219 |
| 7713  | AAAAGATGTTTCATAGCTCTCCTTCACAGCGGGTGTGTTATACCGCCCGAGTTGGCCAGAG   | 7772  |
| 73220 | AAAAGCATTTCTGACAGCGCATGACGCGGGTGAGCTGCTCCACGCCATATGTAAGC        | 73279 |
| 7773  | TTCTGCACGCTCTACCGGAATTAAGGCCAAGTGTTCGGGGCGTGGCGTCTCTCA          | 7832  |
| 73280 | TACTACAGAAATATACACTGAAATGAGGCCAAATGCTGGGGCTCTGGAGGCGAATCG       | 73339 |
| 7833  | TATGTTGACGGCGCTCATCATGGCTATGCGCACTCGGTCTCTGCTACTTATATAATCCC     | 7892  |
| 73340 | TCTGCAATCGAGGCCATTAATGATTAATACTCTTCTTCTGTAAGCTCTACAAACGT        | 73399 |
| 7893  | TGATGCGCAGGGAATACTGAACTTAGACTCAAGACACTGTGAAACTTGTGTTTTCAAC      | 7952  |
| 73400 | ACGATACCGCGGGCTCTCTGAGCAGTCCAGGGCCCTCTGCTTTTGGTTTTTCCAAAC       | 73459 |
| 7953  | CCATCTTTTCAAGAACTTAGCCCCCTATGCACTACTGGCCAAACGGGAAGATGGTTC       | 8012  |
| 73460 | CGGTCTACTTCCGNGGATTAATGGCGCGCTGGAGATCATGACCAAGGGTCAGCTCGCC      | 73519 |
| 8013  | CTGACAAATTAATTTACCAATCACCGGTACGGCCGAGAGAGGGCCCTGTCTGTGACTGGAA   | 8072  |
| 73520 | CTGAAAACTTTTACAGCATCACCGGTTCTGTGTGAAACGCGCGGCCAATATACCAACCGGA   | 73579 |
| 8073  | GTACTAGCGGGATGACGTGTCGGGAACAGCCTTGTCCCGATTTCTTAATCTCTGCCAG      | 8132  |
| 73580 | AGGTCACTGACTGTCTTATCAGGAAGCGTCTCATGCCAGAACTTTAATTTTGGCAA        | 73639 |
| 8133  | TATGGAACCGGGGCTTCTCCCGGCACTTGGTTGACTCACTAGTAATGTCTTAGAAAAATC    | 8192  |
| 73640 | TCTTGGACCGAGGACTGTGTGCGGCTTCCATGTTAGACCTCATGCGATGTGCTGGCAAAAC   | 73699 |
| 8193  | CAGAAATCATCTCTAGCGCCCAACCCCTGAGTCAATTTGTCTATCACAACACGCAACCCCA   | 8252  |



Db 73700 CGCCGCTTATCTGAGCGCCCTGCGCTGAGCGAGTTGTCTATTAGCAAAACCCCATCCCA 73759  
QY 8253 GTCTGCTCAGTCAGTCAGCATTATTACCCCAACCCAGGGCGTTGTCTCCGSCCAATGTT 8312  
Db 73760 ACATGCCGCACACCGTCAGCATCATCCCTTTAAAC--CATCGGTACAGACCCGGCGT 73816  
QY 8313 TTATGACACCTGGAAAGCGGTGTCTACAGAGCATTCACACACGACGAC 8364  
Db 73817 TTATTAGTAGTGGCGAGCGCGGTCTACAGAAATATGTGTACAAACATCCAC 73868

RESULT 6  
AR065850  
LOCUS AR065850 35100 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 18 from patent US 5849564.  
ACCESSION AR065850  
VERSION AR065850.1 GI:5996066  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS Chang, Y., Bohenzky, R.A., Russo, J.J., Edelman, I.S. and Moore, P.S.  
TITLE Polypeptides from Kaposi's sarcoma-associated herpesvirus, DNA encoding same and uses thereof  
JOURNAL Patent: US 5849564-A 18 15-DEC-1998;  
FEATURES Location/Qualifiers  
source 1..35100  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 11.0%; Score 1002; DB 6; Length 35100;  
Best Local Similarity 54.5%; Pred. No. 2.1e-253;  
Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;  
QY 1 ATGGAGAGTTCCGTGCGATGGACCAACACAGTCGACCAAAATCCGGGGTTCATTGTAAC 60  
Db 29792 ATGGACAGTCGGAAGGGTGCACTGATATGGACGAGCCAGCCCGGGTTCTACTCAAC 29851  
QY 61 ATGACGTCCGATCCCAAAGTCAGGGGTGTCTGGATCAGTCAGTCGCTGTCATATA 120  
Db 29852 ATGACATCGATCCCAAAGTCAGGAGCGTAGTGGAAACGATTGACAGGTTGTCAAATATC 29911  
QY 121 ACTACACAGCCACCGGAATGGGTGTGATGACGCTCGCCCTCGATCCGGCTGAAGACTCC 180  
Db 29912 ACTACCTGCCCCGAGATGGCTGGTATGACCTAGAGTTGATCCACTGGAAGACGAA 29971  
QY 181 GGGCCGTTCCTGCGTTTACCGTTTATCTAATTAACGGAACTGTGCTGCTGGGAAAGT 240  
Db 29972 GGGCCCTTTCTGCGTTTTCGGCATACGTAATAACCGGAGCTGACAGGCGGGGAAAGC 30031  
QY 241 ACCAGCATATCGGCCCTGTACCAAAATTTAACTGCTGATCAAGCGCGCGACCCACATA 300  
Db 30032 ACCAGCGTATCCGCCCTTACATCAAGATCTCACTGCTTAATTACGGGGCTACAGTGTA 30091  
QY 301 GCGGCACAGAACCTATCGGTGCGCTTAAGAGAGTTCTGTCCCAAGATCTTCAGCGCTTTT 360  
Db 30092 GCGGCACAGAACTTTCCAGGGCTTTAAAGTCTTACTGTCCCACTATATACCAAGCTTC 30151  
QY 361 GGGTTTAAGAGCGGACATATATAGCGTCAGAAAGCTCATCAGACCGGAGCGTA 420  
Db 30152 GGATTCAAGAGCAGACACATTAATATCTGCCAGGAAGTGCCTCAAGGTAACTCAGTCC 30211  
QY 421 TCCATAGACAAATTCAGCAACAGGAGCTATCGAAGTATTGGCCGGTTATAGTGGACATT 480  
Db 30212 TCCATCGAGCACTCCAGAGATACGAGCTGGCTAGGTACTGGCCAACTGTCCCGATATT 30271  
QY 481 ATGAAGAGGTTATGGGAAACCAATGGCATGTACGGACTATATCCAAACGGGAT 540  
Db 30272 ATTTCGAAATTTATGCGCAAGAAACAAAGGGGCGAGTAGTCTCTCTCAAAGCGCT 30331  
QY 541 TTTGAAACCTCTCGAGAAATGACCGGACCGTGTATTGAGCTTCCAAATATTATTGTAATC 600

Db 30332 TTCAGACTCTTTTCCGCTATGGGTGGAGCCAAATTTGGGACGAGTAACATTATTCGTGATA 30391  
QY 601 GACGAGCCGGAACCCCTGTCTCTTATCATACTCACCACCGCTGTTGTTCTTTTACTGGTTC 660  
Db 30392 GACGAAGCTGGAAACCCCTCTGTCCCATATTTTGAAGCGCGTGTGTTCTTCTTATGTTT 30451  
QY 661 CTAAACAGCTGGCTAAATATACCCCTCTTTACCGCAGGGGGCGGTTCCGTGCATAGTATGC 720  
Db 30452 TACAACAGTTGGGTGGACACCCCGCTATACAGAAATGGTCCCGTGCCTTGCATAGTCTGC 30511  
QY 721 GTCGGTTTCGCAACGACGACAGAAACGCGTTCAGTCAACTTTACAAACACGAGAGCAAAAG 780  
Db 30512 GTGGGGTCTCCACCCAGACGGAGCGCTTTTCAGTCGGTCTTCAACACACGACGACGAGA 30571  
QY 781 ACGGAGATATCGTCTGCGGAGACATCTTAACATTCTATGATAGGAGAGAGGTGATATCT 840  
Db 30572 AACGAGATATCTGCTGTGATAATGTCTCACCTTCTTATTTGGGAAAACGTGAGGTTGCA 30631  
QY 841 GAGTACGTACACTTTGGAGAGAAACTGGGCGCTGTTTATAAACAACAAGCGCTGCACCGAT 900  
Db 30632 GATTATATTAGCTGGACGAGAAATGGGCCCTATTATTAACAATAAGCGCTGACCGAT 30691  
QY 901 CTGAGTTCGAGACCTTCTAAAAATTTTAGAGTATTAATCTTCCCATCTCTGAGCAAGTC 960  
Db 30692 CCCAGTTTGGTCACTTGTCTGAGACCTTAGAATATAATCTAGACATATCACCAGAGTTA 30751  
QY 961 ATGAGTTACGTAGACAGATTGTGTTCTTAAAGTAAGATTATGATCTCTTTAGATAATC 1020  
Db 30752 ATGACTATATAGATAGTTTGTGTTCCGAGAGTAAAGATTCTGGACCGCTCGAGTAT 30811  
QY 1021 ATTGGCTGGACCCGACTCTTTTGTCAATAGCGAGTAAAGCGTATCTTAACAAACCTA 1080  
Db 30812 GCAGGGTGGACAAGACTCTTCATCTCACACGAGGCTGAAGTCTTTTCTGGCAACGCTG 30871  
QY 1081 CACATGCTTAAGCTAGGGGGGATACGAGGACGACGAGCTCTTTTACTCTGTCCTGCTG 1140  
Db 30872 CACACCTGCTGTCGAGTAATAAGGATGCTGTGTCCCAAGCTTTTTCACCTGCGCCAGTG 30931  
QY 1141 GTGTGCGAGGTGTTGTGAAGCCGTTTGAAGATACAAAGCGGCGCTCAACCTCACCAAC 1200  
Db 30932 GTCTGTGAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGGGCGTAGGCTCACACAC 30991  
QY 1201 CTCACCGTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTTCCAGTTT 1260  
Db 30992 ATGACTCCCATAGAAATGGGTAAACAAAATCTTTTTCAGGGCTAAGTAACTACTCGCAGTT 31051  
QY 1261 GTGACGAGGACATGTCATAGTCGCCACGGAATCCACTGAACTGACCCAGCGGTTTACC 1320  
Db 31052 GCTGATCAGGACATGGCTGTGGTGGACCTATATCAAGACGCGTCCACACATCACC 31111  
QY 1321 TTTATCAACAAAGTTTGTAAACACAGCCAGCTGTCTTTTAAATGGAAGAAAACAAAATATGC 1380  
Db 31112 TTCGCCACTAAATTTGTCAAAAACAGCTATGCTACCTACCTTACTTGAAGAGACCAAAAATGT 31171  
QY 1381 ATATGTGGTTTCAGGGTACGTACTTTGAGTTCAAAAGAAATCCTAGACAGCGAACTCTTC 1440  
Db 31172 ATATGCGGGTTTCAGGGTACATCAAAAGATTCAAGTCCATCTTAGACGGGAGCTATTT 31231  
QY 1441 GTGAGAGCGCAATTCGCAAGACCGCTCCGGAATACGTATACCGTCTTTTAAACACGCTACTG 1500  
Db 31232 ATCGAAAGTCATTTCGACGATAACCCCGCTTATGTGTACAGTTTCTTAGTACCGCTGCTA 31291  
QY 1501 TACAAACGCTATGTTCTGTTTACGCGTACGCGCTGCTAGGTACACGAGAAATATCTG 1560  
Db 31292 TATAATGCATGTACTCATTTTACGCGACCGGGGTGAAGCAGGGGCTAGAAAGATTCTCTC 31351  
QY 1561 CAAGACCTTAAATTTGCAACCCCTCCCGGCCCTCTGGCAACCGGGCGGTGTAGACTTTCAA 1620  
Db 31352 AGGACCTCAGGAACTGCCCGGTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 31411  
QY 1621 ACGGTTCTGTAAGAGTTAAACCTGGAAGACGACATCTTCTTACCACGCTGTGTCAGTCCCGC 1680





3597 ACTGCGGCGGAGGAGAACCGGTTGCGCGTGGGGGTAGCCAGCGGGAGAGCGGATGGCGG 3656  
 33626 ACCGAGGCGATCGCGAGTGTGCGGCCAACAGAGAG---GCCAGCGGAGATGGATCTGGG 33682  
 3657 CTTGAGGACGAGATGCTCTGTCGCTTATCAGGTCGCGCTTGTCTATAGCCTTTCTTC 3716  
 33683 CCGTAAGCACACAGGTTGTTCTGTGCGTTTATGAGCGGAGTCTGTCAATGGCTTGGCC 33742  
 3717 CCCACAGCATAAAAACGACCTCTGAAGCTTTGGCAGAAAGGACGCTTATTTATGTAGCTGG 3776  
 33743 CCCACAGGAGAAAAACGCAATGTTCTTAACCTTTGAGGATATGCTACTGATGAACCTCG 33802  
 3777 TGAACCAATCCGACGACCAAGTTGGCGTGTGACCCGCGCTTCCCTCTTCCACCGCTCAGAA 3836  
 33803 TGAACCAATCCGACGACCAAGTTCCCTGTTGAGCGCGCCCTCCCTCTTCCACCGTCAAAA 33862  
 3837 TGGTGTTTAAAGCAGAACTCCCGCTTTTGGCCACAGCTCTAAAGCACCCTGAGAAAGGAG 3896  
 33863 CTGTGTTTAGTAGCAACACACCTTGGCAGGCCAGCTGTGAGGCAACCGCTGGGAGGAG 33922  
 3897 CGGTGAACCCGAGGTACAGTATGCTATCTCTTTAAATAATTTTGGAGCTTTGGAGGTA 3956  
 33923 TACTGAAATTTGGGACGGAAGCTTCTAGCTCTCTAAAGATGCTTCTCAAACTGGGTGGA 33982  
 3957 TACTGTAGTCCGGAGCAACGCTAAAGGCTAGTCCGGTGGCTTGACCGCATGTTAAGGAT 4016  
 33983 CCGTGAATTTGGGATCCACACTAAACGCGGAGGCTAGCTTGGCCCTTGTGCTAGGCT 34042  
 4017 CCGTCCCAAGGATCAGGCTTTATATCTCCGCTTGCACACATAGGACCAACCATAA 4076  
 34043 CCGTCCCTTAGATCAGCACTTTATATCTCTGATCGGACGAGTGGGACCAACCATCA 34102  
 4077 TCTATCTATCGTGGATTAATACAGGTTAGTGTCTATATCCATACCGCTTCAGCA 4136  
 34103 GCTTGTCTGTGGGGATACACTGTGGTGTAGCTTCCGAACTGTCTGAGCA 34162  
 4137 GAGCGCGAGCTCTCTGTTTAAATAATGAGAGATTTAAATAATCAAGCA 4188  
 34163 GCGAGAGCGAGTTCTGTTTCAGAAATGATGAGGCTTCAGAAAGGAATCCA 34214

RESULT 7  
 AR127848  
 LOCUS 35100 bp DNA linear PAT 16-MAY-2001  
 DEFINITION Sequence 18 from patent US 6183751.  
 ACCESSION AR127848  
 VERSION AR127848.1 GI:14115510  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 35100)  
 Chang,Y., Bohenzky,R.A., Russo,J.J., Edelman,I.S. and Moore,P.S.  
 TITLE Unique associated Rapsin's Sarcoma virus sequences and uses thereof  
 JOURNAL Patent: US 6183751-A 18 06-FEB-2001;  
 FEATURES  
 source  
 1. .35100  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

121 ACTACGAGCCACCGGAAATGGGTTGTGATGAGCTGGCTTCATCCGGCTGAGACTCC 180  
 29912 ACTACCTCGCCCGGAGATGGGCTGATATGACTAGATTCGATCCACTGGAGACGAA 29971  
 181 GGGCGGTTCTTGGCGTTTACCGTTTATCTAATTAACGGAACTCTGTGTCTGGGAAAGT 240  
 29972 GGGCGCTTCTTGGCGTTTCTGGCATACGTAATTAACGGGAGCTGACGAGAGCGGAAAGC 30031  
 241 ACCAGCATATCGGCGCTGTACCAAAATTTAAATCTGCTGATACCGGGCGGACCACTA 300  
 30032 ACCAGGTTATCGGCGCTTACATCAGATCTCACTGCTTAATTAAGGGGGCTACAGTGGTA 30091  
 301 GCGGACAGAACTATCGGCTGCGCTAAAGAGCTTCTGTCCACGATTTTACGGCTTTT 360  
 30092 GCGGACAGAACTTTTCCAGGGCTTTTAAAGTCTTACTGTCTCCACTATATACCAAGCGCTTC 30151  
 361 GGTCTTAAAGAGCGCACATCAATATAGCGCTCAGAAAGCTCATCAGACCGGAGCGTA 420  
 30152 GGAATCAAGAGCAGACATTTATCTGCGAGGNAAGTGCCTCAAGGTAATCTCAGTCC 30211  
 421 TCCATAGAGCAATTTAGCAACAGGAGCTATCGAAGTATTTGGCGGGTTATAGTGGACATT 480  
 30212 TCCATAGAGCAATTTCCAGAGATACGAGCTGGCTAGTGTAGTGTGGGCAACTGTCAACGATATT 30271  
 481 ATGAAAGAGAGGTTATGGCGGAAAAAACCCTATGGCATGTACGGGACTATATCAACGCGAAT 540  
 30272 ATTGAGGATTTATGCGCAGAGAAACAAAGGGGAGTATAGCTCCCTCTCTCAAGCGCT 30331  
 541 TTGGAACCTCTCGAAGATGACCGGACCGTGTATTTGAGACTTCCAAATATTATTTGTAATC 600  
 30332 TTGAGACTCTCTTGGCGTATTTGGGTGGAGCAATTTGTGGACGAGTAACATTTATCGTGATA 30391  
 601 GACGAGCGGAAACCCCTGCTCTTACATACTCACCAACCGTGTCTGTTCTTTTACTCGGTC 660  
 30392 GACGAGCTGGAACCTCTCTGCTCCATATTTTGAAGCGCTGTGTTCTTCTATTGTTT 30451  
 661 CTAAACAGCTGGCTTAATACCTCTTTTACCGGACCGGGCGGTTCCGTGCTATAGTATGC 720  
 30452 TACAACAGTGTGGCTGGACACCCCGCTATACAGAAATGGTGGCGCTTGTGCTATAGTCTGC 30511  
 721 GTCGGTTTCGCAACGAGACAAACCGGTTCCAGTCAACTTACAACCAACGCGGACGCAAAAG 780  
 30512 GTGGGCTCTCCACCCAGAGCGGCTTTCAGTCCGCTCTTCAACCAACGCGGACGAGAGA 30571  
 781 ACGGAGATCTGCTGCGGAGAACATCTTAAATTTAGAGTATTAATCTTCCCATTCCTGACGAGTC 840  
 30572 AACGAGATATCTGCTGTGATAATGTGCTCACCTTCTTATTTGGGAAACGCTGAGGTTGCA 30631  
 841 GAGTACGTACACTTGGAGAGAACTGGGCGCTGTTTATATAAACAACGCGCTGACCGAT 900  
 30632 GATTATATTAGGCTGGAGAGATTTGGGCTTATTTATATAAACAATAGCGCTGTACGGAT 30691  
 901 CTGCAAGTTGCGACACCTTCTTAAATTTTAGAGTATTAATCTTCCCATTCCTGACGAGTC 960  
 30692 CCCAGTTTGGTCACTTGTGTAAGACCTTAGAATATAATCTAGACATATCACCAAGATTGA 30751  
 961 ATGAGTTACGTAGACAGATTTGCTTCTTAAAGTAAGATTATGATTCCTTTAGATAATC 1020  
 30752 ATGAGCTATATAGTAGTTTGGTTCCGAGAGATTAAGTTCTGGAACCGCTCGAGTAT 30811  
 1021 ATTGGCTGGACCGGCTCTTTTGTGCATAGGAGGTAAGCGGTATCTAACAACCTTA 1080  
 30812 GCGGGTGGACAGACTCTTCTATCTCACACCGAGGAGTGAAGTCTTTTCTGGCAAGCTG 30871  
 1081 CACACATGTCTAACGCTAGGGGCGGATACCGAGGACAGGAGCTTTTACCTGTCTCCGCTG 1140  
 30872 CACACCTGCTGTGAGTAATAAGATGCTGTGCCAACAAGCTTTTACCTGCCAGTG 30931  
 1141 GTGTGCGAGGTGTTTGTGAGCGGTTTGAAGTAATAAAGGGGCGCTCAACCTCACCAAC 1200  
 30932 GTCTGTGAGGTGTTTACAGAGCCATTTGAGGAGTACAAACGCGGCGGTAGGCTCAGAC 30991  
 1201 CTCACCGTGCAGGTTGGGTGACAAAAACCTCTTTAAGTAAATTTATTCGAGTTT 1260



QY 3211 TGAGTCGTGTTCTCCGAATGAGAGTCGGTTTCGTCGCGCTTTTGGAGTAGGGAG 3270  
 Db 33206 ATTGCGCGCTGACTGATCGGGGATATCTGTAGATCCAGAGGGTGTCTGGCGCGATG 33265  
 QY 3271 GGAATCTTTGGTCAAAAAATCCTCAGGTGCAATCATGTCATCAATCAATCGGGGCTCCGT 3330  
 Db 33266 GCGTGTCTCTCGGGAAGAGCTGTCTGGGGGAGACATATCTATCAGCGTGGTCCAGCAT 33325  
 QY 3331 ACCGGGCGGTACCGGCAACCCGATACGGGA-----CCCGGGAACAAAC 3378  
 Db 33326 AGCCGCGCGCTGCGCAATCTCTGGAAGTGATGAAGAGGTGGAGTGAATATGAAT 33385  
 QY 3379 GTCTTGTCTCTTTTCTTAGTGTCCCGGGAATCGGCAGCATCTCTGGAGTGGCGGG 3438  
 Db 33386 TCACGGGGGTGCTCTCGAGGGCTCTTCAATTGGAAGCATCTCTTCAATCGTGTG 34445  
 QY 3439 GGTCTCTTGAAGGGTCGATGAAC-----ATACCCATGCC 3476  
 Db 33446 TGCTAGACGAGTCTCTCAAAACATCGCATGGCCCTTGACGGGGTGTAGCCCTAGGGG 33505  
 QY 3477 GAAACGTTCACCTGTAAACCGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGG 3536  
 Db 33506 GGAATTTCAAGCACACGAGTTATGCTTTACTGCTCCACAGGCCCTCAGTCCACAG 33565  
 QY 3537 CACCCCTCCGATTTCAACTAGATGATCAATGCAAGTTTAAATGTTGACGCTAGAA 3596  
 Db 33566 TCTCAGCCCGTGGCGAGTCAAAATAGTCTGTGGTAAAGTGAATACAGCCCTGGA 33625  
 QY 3597 ACTGGCGGAGGAGAACCGGTGCGCGGTGGCGGTAGCAGGGGAGAGATGCGGG 3656  
 Db 33626 ACCGAGGCCATCCGAGTGTGGCCACCAAGAGAG--GCCAGCGAGATGATGCTGG 33682  
 QY 3657 CTTTGAGGAGAGATGTCTCTCGCTTTATCAGCGTCCCTTGTCTATAGCCTTTCTTC 3716  
 Db 33683 CCGTAAGCACCAAGTGTCTCTGCTGCTTATGAGCGAGTCTGTCAATGGCCTTGGCC 33742  
 QY 3717 CCCAGGCAATAAAGCGACCTCTGAGCTTGGCAGAAAGGAGTATTTAGTCTGG 3776  
 Db 33743 CCCAGGAGAAAGCGCAATGTTCTAATTTGAGGATGATGCTGATGAAGAACTCG 33802  
 QY 3777 TGAACCAATCCAGGCAAGTGTGGCGTGTGACCCCGCTTCCCTTTCCACCGTCAGAA 3836  
 Db 33803 TGAACCAATCCAGGCAAGTGTGCTGTGAGCGCGCTTCCCTTTCCACCGTCAAAA 3862  
 QY 3837 TGTGTTTAAAGCAGAACTCCCGCTTTGCGCAACAGTCTAAGACCCGTCGAGAGGAG 3896  
 Db 33863 CTGTGTTTAGTACAAACACACCCCTGGCGAGCCAGCTGTGAGGCAACCCGTCGGAAGGAG 33922  
 QY 3897 CGGTGAACCCAGGTACAGTATGGCTATCTCTTTAAATAATTTTGGCGCTTGGAGTGA 3956  
 Db 33923 TACTGAAATTTGGGACGGAAGCTCTAGCTCTCTAAGATGCTTCTCAACTGGTGGAA 33982  
 QY 3957 TACTGTAGTCCGAGCAACGCTAAGGCTAGTCCGCTGGCTGACCGGATGGTAAGGAT 4016  
 Db 33983 CCGTGAATTCGGATCCACACTAAACCCAGGCGAGTAGCTTGGCCCTTGTGTGCGGT 34042  
 QY 4017 CCGTGGCAAGATCAGACTTTAATCTCCCGTTGCGCAACATACGACCCACCATAA 4076  
 Db 34043 CCGTGGCTAAGATCAGACTTTAATCTCTGATCGGAGAGTGGGACCCACCATCA 34102  
 QY 4077 TTCTATCTATCGGTGGAATAATCAGGTTACGTTGTCTCATATCATAACCGCTTCAGCA 4136  
 Db 34103 GCTTGCTCTGTGGGGATACACTGTGTGTGTAGCTTAAGTTCGCCAATCTGTCTGAGCA 34162  
 QY 4137 GAGCCGCGAGCTTCTGTTTAAATTTGGAAGAAATTTAAATTTCAAGCCA 4188  
 Db 34163 GCGAGACAGTCTTCTGTTTTCAGAAATGATGAGAGGCTCAGAAAGGAATCCA 34214

RESULT 8  
 AR194750 35100 bp DNA linear PAT 20-APR-2002  
 LOCUS

DEFINITION Sequence 18 from patent US 6348586.  
 ACCESSION AR194750  
 VERSION AR194750.1 GI:20241342  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 35100)  
 AUTHORS Chang, Y., Bohenzky, R.A., Russo, J.J., Edelman, I.S. and Moore, P.S.  
 TITLE Unique associated Kaposi's sarcoma virus sequences and uses thereof  
 JOURNAL Patent: US 6348586-A 18 19-FEB-2002;  
 FEATURES  
 location/Qualifiers  
 source  
 1..35100  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

Query Match 11.0%; Score 1002; DB 6; Length 35100;  
 Best Local Similarity 54.5%; Pred. No. 2.1e-253;  
 Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;

QY 1 ATGAGAGAGTTCGCTCGGATGACCAACACAGTCGTAACCAATCCGGGTTTCATCTTGAAC 60  
 Db 29792 ATGCACAGCTCGGAAGGTTGCACTGATATGACGAGCCAGCCCGGTTTCATCTCAAC 29851

QY 61 ATGACGTCGATGCCAAAGTCAGGGTGTGTGTGATCAGTCAGTCGCTGTCAATATA 120  
 Db 29852 ATGACATCGGATGCCAAAGTCAGGAGCTAGTGAACAGATTGACAGGTTGTCAAATATC 29911

QY 121 ACTACACGCCACCGGAAATGGGTTGTGACACCTGCGCTTCGATCCGGCTGGAAGACTCC 180  
 Db 29912 ACTACCTCGCCCGGAGATGGGCTGTATGACCTAGATTCGATCCACTGGAAGACGAA 29971

QY 181 GGGCGGTTCTTTCGCTTTTACCGTTTATCTAATACGGGAAGTGTGTGTGTTGGGAAAGT 240  
 Db 29972 GGGCGGTTCTTTCGCTTTTCGGCATAGCTAATAACGGGAGTGCAGGAGCGGGGAAAGC 30031

QY 241 ACCAGCATATCGGCGCTGTACCAAAATTTAAACTGCTGTACGCGGCGGACCAACATA 300  
 Db 30032 ACCAGGTTATCGGCGCTTACATCAGATCTCACTGCTTAACTAGCGGGCTACAGTGGTA 30091

QY 301 GCCGACAGAACTTATCGCTCGCTTAAAGACGTTCTGTCCAGCATCTTCAGGCGCTTTT 360  
 Db 30092 GCGGACAGAACTTTCCAGGCTTTAAAGTCTTACTGTCTCCACTATATACAGCGCTTC 30151

QY 361 GCCTTTAAGAGCCGACATCAATATAGCGCTCAGAAAGCTCATCAGCGGAGCGGTA 420  
 Db 30152 GGAATCAAGAGCAGACATTAATATCTGCCAGAGGAAGTGCCCAAGGTAACTCAGTCC 30211

QY 421 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTGGCCGCTTATAGTGGACATT 480  
 Db 30212 TCCATAGAGCAACTCCAGAGATACGAGCTGGCTAGGTACTGGCCAACTGTCAACGATATT 30271

QY 481 ATGAAGAGGTTATGGCGAATAAAACCCATGGCATGTACGGGACTATATATCCACGCGAAT 540  
 Db 30272 ATTCGAAATTTATGCGCAAGAAACAAAGGGGCGAGTATAGTCCCTCTCTCAAGCGCT 30331

QY 541 TTGAAACCCCTCTCGAAGTACCGGACCGGTTTATGGAATTCCTCAATATTATTGTATATC 600  
 Db 30332 TTCAGACTCTTTCGCGTATGGGTGGAGCCAAATTTGTGGACAGTAACTATCTGTGATA 30391

QY 601 GACGAGCGCGAAACCTGTCTCTTACATCTACACCGCTCGTGTCTTTTACTGTGTTTC 660  
 Db 30392 GACGAAGCTGGAACCCCTCTCTGCTCCCATATTTTGACGCGCGTGTCTTCTTATTTGGTTT 30451

QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTTACCGCAGGGGGGGTTCGGTGCATAGTAGTC 720  
 Db 30452 TACAAAGTGGCTGGACACCCCGCTATACAGAAATGGTGGCGCTTGTGATAGTCTGC 30511

QY 721 GTCGGTTCCCAACGACAGCAAAACCGGTTCCAGTCAACTTACAAACGAGGACGCAAAAG 780  
 Db 30512 GTGGGCTCTCCACCCAGACGACGCTTTTCAGTCTGGTCTTCAACACACGACGACGAGA 30571

Qy 781 ACGGAGATATCGTGTGCGGAAACATCTTAACATTCATGATAGGGAAGAGTCTGATCT 840  
 Db 30572 AACGAGATATCTGCTGTGATATATGCTCACCTTCTATTGGGAAACGAGGTTGCA 30631  
 Qy 841 GAGTACGTACACTTGGAGAGAACTGGGCGCTCTTATTAACACAGAGCGCTGCCAGT 900  
 Db 30632 GATTATATTAGGCTGGACAGAAATGGGCCCTATTATTAACAAATAGCGCTGTACGAT 30691  
 Qy 901 CTGCACTTCGACACACCTTCTAAATAATTTAGAGTATAATCTTCCCATTCCTGACGAAGTC 960  
 Db 30692 CCCAGTTTGGTCACTTGTCTGAAGACCTTAGAATATAATCTAGACATATACACAGAGTTA 30751  
 Qy 961 ATGAGTTACGTAGACAGATTTGCTGCTTAAAGTAAAGTATAGATCTTTAGATATAC 1020  
 Db 30752 ATGAGCTATATAGATAGTTTGGTTTCCGAGAGTAAAGTATTCGACCCGCTCGAGTAT 30811  
 Qy 1021 ATTGGCTGGACCCGACTCTTTTGTACATAGCGAGTAAAGCGTATCTAAACAACTTA 1080  
 Db 30812 GCAGGTTGACAGACTCTTCACTCACACGAGAGTGAAGTCTTTCTGGAACGCTG 30871  
 Qy 1081 CACATATCTAAGCTAGGGGCGGATACACAGGACACAGAGTCTTTTACCTGTCCTG 1140  
 Db 30872 CACACCTGCTGTGAGTAAAGGATGCTGTGTCCACAAAGCTTTTCACTGCCCAGTG 30931  
 Qy 1141 GTGTGCGAGTGTGTTGAGAGCGTTTGGAGGATACAGAGGCGCGTCAACCTCACCAAC 1200  
 Db 30932 GTCTGTGAGTGTGTTACAGAGCCATTTGAGAGTACAAACGCGCGGTAGGCTTCAACAC 30991  
 Qy 1201 CTCACCTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTTCGAGTTT 1260  
 Db 30992 ATGACTCCCATAGAAATGGGTAAACAAAATCTTTTTCAGGCTAAGTAACTACTCGCAGTTT 31051  
 Qy 1261 GTGACACGACATGTCATAGTCGCCACGGAATCCACTGAAGTTTCGACCCAGGTTACC 1320  
 Db 31052 GCTGATCAGACATGGTGTGTTGGGACCTTATACACAGCGGTCCACAGATCAC 31111  
 Qy 1321 TTTATACCAAGTTTGTAAACACAGCCAGCTGTCTTTAAATGAAACAAACAAATGC 1380  
 Db 31112 TTGCGCACTAAATTTGTCAAAACAGACTAGCTAGCTTACTGAAAGACCAAAATGT 31171  
 Qy 1381 ATATGCGGTTTCAGGTAGCTACTTTGAGTTCAAGATCTTAGACGCACTCTTC 1440  
 Db 31172 ATATGCGGTTTTCAGGTTTCAACCAAGATTCAAGTCCATCTTAGACGCGGAGCTATT 31231  
 Qy 1441 GTGAGACGCAATTCGCAAGACCGTCCGGAATAGTATACGGTTTCTTTAAACACGCTACTG 1500  
 Db 31232 ATCGAAAGTCAATTCGCAAGTAAACCCCGCTTATGTGTACAGTTTCTTAGTACCTGCTA 31291  
 Qy 1501 TACAAAGCTATGTAATCTGTTACGCGTACGCGGTGACTAGGTCAACAGGAAATATCTG 1560  
 Db 31292 TATAATGCCATGTACTCAATTTAAGCGCACGCGGTGAAGCAGGCGGATGAAGAAATTCCTC 31351  
 Qy 1561 CAGACCTTAAATTTGCAACCTCCCGCGCTCTGGCAACCGGCGGTGTAGACCTTCAA 1620  
 Db 31352 AGGACCTCAGGAACTGCGGCTGTCTCAAGAGCTGATCTCTGAGATGAGCTCGAGGAC 31411  
 Qy 1621 ACGTTTCGTGAAGATTAAACCTGGAAAGACGACATCTTCTACACGTTGTCAGTCCCGG 1680  
 Db 31412 GTTCTGGGCGAGGAG-----GGGACACAGATGCTCTTCTACCTCACCGCGACCTCCCA 31465  
 Qy 1681 CCACCGCGGATACCTCCCTCCAGGTTTGTGTCACACGACTACTGCGCCCTTAAAGGAC 1740  
 Db 31466 CCATCCCGCCACCGCGGCTCTTCAACACTGTGTGCTTACTTCTCGGGGCGAAGAA 31525  
 Qy 1741 GTGTTGCGCTCCAGAAATAAGAGTGGCGTGTGCTGTTTGGCGGGGAGTTTGAGAAAGAA 1800  
 Db 31526 CTATTTGCAACAGGCTGGCCCTGGCACGCGGACACTTTGGTACAGGTTCTCTCCACTCC 31585  
 Qy 1801 ACGTTTCCGGTTTACGGTTTAAACATGTTGTAAGGACGAGTGTGACTTTGCTCCCT 1860  
 Db 31586 GATTTTTACAGTTTACGGTGAACATCTGTGTGCGAGATGCGGTGGACTTTGTGTCACT 31645  
 Qy 1861 TCAGAACCTCTCAACGGGCTGTGGCGTTTGCATCGACCGTTGAATCGTATATAAATTAAG 1920

Db 31646 TCCCGCGGCTCCACGGTCTAGTGGCATACGATCCACTATAGACACTATATATCCAG 31705  
 Qy 1921 GGTACACGTTTACCGGTAGCGTTCGGTCTGCTCAGGCTTTGCCACTCAGTGTATGAC 1980  
 Db 31706 GATATAGTTCCTCCAGTCAGATTGGCGCTCCAGGAGGACAGCGCTCAGCGAGAC 31765  
 Qy 1981 CTCAGGAAGAAGATGCGCTCCTCGTGTGTCAGACTCTAGCGGTTTATCGCTGCTA 2040  
 Db 31766 CTGCGCAAAAAGATGCGCTCCTCATAGTTGTCAGGACTCATCGGGTTCATTGCTGCTGCTG 31825  
 Qy 2041 GAGATATACATAAACAATTCAGCAACCATGAGGACGGGAGGTTTCCAAAGTGTGC 2100  
 Db 31826 GAAATAACGTCACAGATACAGAGACCTCGAAGTGGCGAGCTGTTTAAACATATGT 31885  
 Qy 2101 TGTGCGGGGACTATGGGTCAGTCAAAATTTAGCCATGACCATCTGTAAGAGCAGGGA 2160  
 Db 31886 TGTGAGGGGACTACGGTATCAGTCTTAATCTGGCTATGACCATAGTGAAGGACAGGG 31945  
 Qy 2161 ATGTCGTTGAGAGCAGTAGCCGTAGTATTTGGGTCACCAAGAAAGCTCCACAGAGCCAC 2220  
 Db 31946 GTTTCACATAAGTAGGGTGGCCATATCGTTTGGCAACCCACCAATATCAGAGCCAGTCTA 32005  
 Qy 2221 GTGTATGTAGCAATATCAAGGGCTGTTAACTCAAATTTATTTGGTCATGGAAGCAACCCC 2280  
 Db 32006 GTGTATGTGGGTGATCCAGGGCCATCGACGCTCGTTACCTGTGTAATGACAGTAATCCC 32065  
 Qy 2281 CTTAAAAACCTCTCAGAGAACACGTCGATTAACCTCCGCGCAAGCATATAGTCCGGCC 2340  
 Db 32066 CTTAAGCTAAATGGACCGGGTGAGCCCGAGTCCCACTCTCAAAAGTACATCAAGGCC 32125  
 Qy 2341 CTCCACAAACCCAAACACAAACCTCATCTACTAA----- 2373  
 Db 32126 CTATGCAACCCAGACTACTCTGATCTACTGACCCGTAACCCCTCTCTTAGGACACTGAT 32185  
 Qy 2374 -----ATAATAAATAAACAAGCAANTGAAACATTTGTG 2406  
 Db 32186 GTGTTTGGGAATAAAGCATAGACTTTGACACTATAATGTTCTGTATTGACACCATTTCT 32245  
 Qy 2407 TTTTATTTCAGTCCAAACACCGCCAGGATAGTTGTCATTTTCCACACACCGGGGGGTGG 2466  
 Db 32246 TTAATTATCAGTCCAGCCACCGCCAGTTATGACCGTTTCCACACAGGGGTGCGTGG 32305  
 Qy 2467 CGCGCATAGTTTGTAGCAGCAAGAACCGCCGGGCGCTCTGGGTTAGAGTACGCC----- 2521  
 Db 32306 AGGCCAGGATGCGGGTTGGTGTGACCTGACCTGACCCCGGTAGTTGTGTTCTCTGATG 32365  
 Qy 2522 -----GCCCCCGTGTGAGGTCATACGTCCTTTAGGGCCCTGGGTT 2564  
 Db 32366 AAATCGAGTGGCGGAAGTACTGGGAGATTGGTTGGAGGTGACCTTTGTCTCGAGC 32425  
 Qy 2565 GTACCAACATATAATCAGCTCGTCATCCGTTGAGTCCCTGCGTGAAGCGGCGCTCTTG 2624  
 Db 32426 GAGACACGATCAGCTCACGCGGACGAGGGCTCTCTGTGTGTCTCTCCCGGAGATA 32485  
 Qy 2625 CAGCGCGCACGAAACACCGCGCTTTGTAACGTTGATCT-----GCAT 2664  
 Db 32486 TAAATATCAGGACGCCACTGCTTTGCGCTTAAAGTTTGGTTGTCTCTGCGAGCGCACCA 32545  
 Qy 2665 CGTCTGCTGTCTGAAGATGAACACAGGTCAATCAATCAGTGGCGCGCGCTGCTCAT 2724  
 Db 32546 CATCTCGTACACAGAGGCGGTGAGACTGCTTTTGGCTTCTGGGCCACGTCATGA 32605  
 Qy 2725 CGGCGTGTATTCGTCCTCATGTCACAGAGCTCGTCTGCTGCTGCTGCTCATCTG 2784  
 Db 32606 GCCCGATTCTCTGATCAATACTCTCCCTTGGTCTTCTCCGCTCTCTCGGACGAGGGTG 32665  
 Qy 2785 AAATAATAATGCTCTTCTGCTGTCAGTTGATCCAGCAACACAGGTTCTTTCTCATCGG 2844  
 Db 32666 GCTGTGTGGGAAAAATGGCGCGCTCGGTAAACGGGCGCTCATTTGTTACGCTCCGAGAGT 32725  
 Qy 2845 AAGACATCCCGCCCTCTCGCGGTGATCTAGATATATCACTGACACAGCGCGGCGAGGCC 2904

```
Db 32726 TGAACCTGTCATCGCTATCAGATCCGATGTCAGGTCGACGATCGCGTGGTGCAGCGC 32785
QY 2905 G----- 2905
Db 32786 GCAGGGGGCCACAGAGGCCCTTCATCAGGTCGCTGTATGTGTAACCTTTGTTGCCAG 32845
QY 2906 -----TGCGAAAGCGCTCAGCTTTCGCGCACCTTCATCAGTGT 2946
Db 32846 GTACACTATTTCGGAAGCAGGTGAAGTCCGTATGCCCGGTCCAGTGTATGCGGCCA 32905
QY 2947 CCTCGTAGTCATCTATAAATAATCGCTTCGCTGTCTGATTAGATGACTGGGTGCGC 3006
Db 32906 TCGGTTCCAGATAGAACCCCTTCGCTGCTGAAGGTGAGAGCCAGCAGGGGAAATC 32965
QY 3007 GGTGTTGGCCAGTACCAACCGCAACGGGC-----CGCGACCAACAAACGCG 3053
Db 32966 CGTCATCCTGACTAACCCATCCCATGACGCTCGGACTCGCGTGTGTTGAACGTC 33025
QY 3054 GTCACTGGCGGGGGCGCTCGGTGTGTGGTGGCACCTGTGAGTAATGACGGTAGGT 3113
Db 33026 GCACGCGCGCGCTACCACTGCTACCGGTTTGGGCGTATGGGCGGCTCTGGCCAGAGGC 33085
QY 3114 AGCGCGCGCTCTGTTAAAG-----ACTCACCGCTCGCGT 3150
Db 33086 TCGGGCGCAGTGAGATAGGTTGAAAAAGTCTGCAGGCTACCCCTCTGGCTCGCTT 33145
QY 3151 CGTCGGATCTATTCTAGTGGGAGGAGACCGCGGTTCTGTCTTCTTATCCACATCAC 3210
Db 33146 CCTCCTGAACATCGCTCATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 33205
QY 3211 TGTAAGTCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3270
Db 33206 ATTCGCGCTCGACTGATCCGGGATATCTGATATCAGAGGGTTGCTGGCGCGATG 33265
QY 3271 GGGAACTTTGTTCAAAACATCTTCAGTGCAGTCAATCACTTGGGGGCTCCCGT 3330
Db 33266 GCGGTCTCTCGCGAGACGCTGCTGGGAGACATATCTATCACGCTGGTCCAGCAT 33325
QY 3331 ACCGGGGGTACCGGCAACCCCTAGCGA-----CCGGCGCAACAAAAC 3378
Db 33326 AGCGCGCGGCTGCCAATCTCTGGAAGTATGAAGAGTGGAGTGGGAATATGAAT 33385
QY 3379 GTCTGTGTTCTTTTCTAGTCTCCCGGGAATCGGCAGATCTCTGGAGTGGCGGCG 3438
Db 33386 TCACGGGGGTGCTGTCGGAGCGCTCTTCAATTTGAAGCATCTCTTCTATCGTGTG 33445
QY 3439 GGGTCTTGACGGTTCGGATGAAGAAC-----ATAGCCATGGCC 3476
Db 33446 TGCTAGACGAGTCTCTCAAAACATCGCCATGGCTTGTACGGGGTTGACCGCTAGGGC 33505
QY 3477 GAAAGTTCACCTGTAAACCGCACTGCTATCCCGATCAGTTAATTTCCAGTCAACGG 3536
Db 33506 GGAATTTTAAAGCAGCAGGTTATTGTCCTTTACTGCTCAACAGGCCCAATCCACAG 33565
QY 3537 CACCCCTCCGATTCGAACTAGATAGTCAATTCGAAGTTTAAATGGTTGACCCCTAGAA 3596
Db 33566 TCTCACGCGGTGGCGAGTCAATAGTCTGTTGGCTAGTTAAAGTATTACAGCCCTGGA 33625
QY 3597 ACTCGGCGAGGAGAACCGGTTCCCGGTGGCGGTAGCAGGGGAGACGGATGGCG 3656
Db 33626 ACCAGGCGCATCGCGAGTGTGGCGCAACAGAGAG-----GCACGGGAGATGGATGTGG 33682
QY 3657 CTTTGAGGACGAGATGCTCTGTCGCTTTATCAGCGTCGCTGTCTATAGCCCTTTCTTC 3716
Db 33683 CCGTAAGCACGAGGTGTTCTGTCGTTTATGAGCGGAGTCTGTCAATGGCTTGGCC 33742
QY 3717 CCGCAGCATAAACCGCACTTGAAGCTTGGCAGAAAGGACCTTATTATGTAGTGG 3776
Db 33743 CCGCAGGAGAAAAACGCAATGTTCTAACTTTGAGGATATGCTACTGATGATGAACCTCG 33802
QY 3777 TGAACCAATCCAGCAGGTTGCGTGTGACCCCGCTTCCCTCTTTCACCGCTCAGAA 3836
Db 33803 TGAACCAATCCAGCAGGTTGCGTGTGAGCCCGCTTCCCTCTTTCACCGCTCAGAA 3882
```

---

```
QY 3837 TGGTGTTTAAAGCAGAACTCCCGTTCCTGCTTTGCCCACAGCTTAAGCACCCGTCGAGAAGGAG 3896
Db 33863 CTGTGTTTAGTAGCACAACACCTCGCGAGCCACGCTGTCGAGGACCCGTCGGAAGGAG 33922
QY 3897 CGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAAAATATTTTGGGCTTGGAGGTA 3956
Db 33923 TACTGAAATTTGGGAGCGGAAGCCTTAGCTCTCTTAAAGATGCTTCTCAAACTGGGTGAA 33982
QY 3957 TACTGTAGTCCGAGCAACGCTAAAGGTAGTCCGGTGGCTTGACCGGATGGTAAGGAT 4016
Db 33983 CCTGACATTTGGGATCCACACTTAACGCGGCAAGTGTGGCCCTTGTGTACGGT 34042
QY 4017 CTTGCCCAAGGATCAGACTTTTAATATCTCTCGGTTCGGAACAATACAGCACCACTAA 4076
Db 34043 CTTGCCCTAAGATCACCCTTTTAATATCTCTGATCGCAGCAGTGGGACCACTCA 34102
QY 4077 TTCTATCTATCGTGGATTAATCAGGTTAGTTCCTCATATCCATAACGCGCTTCAGCA 4136
Db 34103 GCTTGTCTGTGGGATACACTGTGGTGTAGCTTAAGTCCCGATCTGTCTGAGCA 34162
QY 4137 GAGCGCGGAGCTTCTGTTTAAATGAGAAAGATTTTAAAAATTCAGGCCA 4188
Db 34163 GCGAGAGCAGTTTCTGTTTTCAGAAATGATGAGAGGCTCAGAAAGGAATCCA 34214
```

---

```
RESULT 9
BD190715 35100 bp DNA linear PAT 17-JUL-2003
LOCUS Unique associated Kaposi's sarcoma virus sequences and uses
DEFINITION thereof..
ACCESSION BD190715
VERSION BD190715.1 GI:33000454
KEYWORDS JP 2002513274-A/15.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 35100)
AUTHORS Chang, Y., Bohenzky, R.A., Russo, J.J., Edelman, I.S. and Moor, P.S.
TITLE Unique associated Kaposi's sarcoma virus sequences and uses thereof
JOURNAL Patent: JP 2002513274-A 15 08-MAY-2002;
        THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK
COMMENT PN JP 2002513274-A/15
        PD 08-MAY-2002
        PF 22-JUL-1997 JP 1998509105
        PR 25-JUL-1996 US 08/686349,25-JUL-1996 US 08/687253 PR
        25-JUL-1996 US 08/686350,25-JUL-1996 US 08/688814 PR
        25-JUL-1996 US 08/686243,05-SEP-1996 US 08/708678 PR
        10-OCT-1996 US 08/728323,13-NOV-1996 US 08/748640 PR
        13-NOV-1996 US 08/747887,29-NOV-1996 US 08/757669 PI YUAN
        CHANG, ROY A BOHENZKY, JAMES J RUSSO, ISIDORE S EDELMAN, PI PATRICK
        S MOORE
        PC C07H21/04,C12Q1/68,C12P19/34,C12N15/10
        CC Strandedness: Double;
        CC Topology: Linear;
        FH Key Location/Qualifiers.
FEATURES             Location/Qualifiers
     source            1..35100
                        /organism="Homo sapiens"
                        /mol_type="Genomic DNA"
                        /db_xref="taxon:9606"
ORIGIN
Query Match          11.0%; Score 1002; DB 6; Length 35100;
Best Local Similarity 54.5%; Pred. No. 2,1e-253;
Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;
QY 1 ATGGAGAGTTCGTCGGATGGACCAACACGTCGACCAATCCGGGTTCTATCTTGAAC 60
Db 29792 ATGGACAGTTCGGAAGGTTGCACTGATATGGACGAGCCCGGGTTCATCTCTAC 29851
QY 61 ATGACGTCGCGATGCCAAAGTCAGGGGTGTCGTGGATCAGTCAGTCGCTGTCATAATA 120
```

|    |       |  |       |  |
|----|-------|--|-------|--|
| Db | 29852 | ATGACATCGGATGCAAGTCAGAGCGTAGTGGACAGATTGACAGGTTGTCAATATC        | 29911 |  |
| Qy | 121   | ACTACAGCCACCGGAAATGGGTGGTACGACCTGGCCCTTCGATCGGCTGAAGACTCC      | 180   |  |
| Db | 29912 | ACTACCTCGCCCGGAGATGGGCTGGTATGACCTAGAGTTTCATCCACTGGGAACGAA      | 29971 |  |
| Qy | 181   | GGGCGGTTCTGGCGTTTACCGTTTATCTAATTACGGGAACTGCTGGTCTGGGAAAGT      | 240   |  |
| Db | 29972 | GGCCCTTTCTGGCGTTTTCGGCATAGTAAATACGGGGACTGACGAGCGGGGAAAGC       | 30031 |  |
| Qy | 241   | ACGAGCATATGGCCCTGTACCAAAATTTAAACTGCTGATACAGGGCGGACCAACATA      | 300   |  |
| Db | 30032 | ACCAGCGTATCGGCCCTACATCAGATCTCAACTGCTTAATACGGGGCTACAGTGGA       | 30091 |  |
| Qy | 301   | GCCGACAGAACCTATCGCGTCGCTAAAGACGTTCTGCCAGATCTTTCAGCGTTTTT       | 360   |  |
| Db | 30092 | CGCGCACAGATCTTTTCAGGGCTTTAAAGTCTACTGTGCTCCATATACACAGGCTTC      | 30151 |  |
| Qy | 361   | GGCTTTAAGACCGACACATCAATATAGCCGTGAGAAAGCTCATAGACCGGAGCGGTA      | 420   |  |
| Db | 30152 | GGATTCAGAGACAGACATTAATCTGTCAGAGGAAAGTGCCCAAGGTAACTCAGTCC       | 30211 |  |
| Qy | 421   | TCCATAGAGCAATTCAGCAACGAGGATCTCGAAGTATTGGCCGGTTATAGTGGACAT      | 480   |  |
| Db | 30212 | TCCATCGAGCAACTCCAGAGATACGAGCTGGCTAGGTACTGGCCAACTGTACCGATAT     | 30271 |  |
| Qy | 481   | ATGAAAGAGGTTATGGCGAAATAACCCCAATGGCATGTACGGGACTATATCCACGCGAAT   | 540   |  |
| Db | 30272 | ATTCGAGAAATTTATGGCAAGAAACAAAGGGCGAGTATAGTCCCTCTCTCAAGCGCT      | 30331 |  |
| Qy | 541   | TTTGAAACCTCTCGAGATGACCGGACCGTGTATTATGGATCTCCATATATTGTATATC     | 600   |  |
| Db | 30332 | TTTCAGACTCTTTCCGCTATGGGTGGAGCCAAATTTGGGACGAGTAACATATTCGTGATA   | 30391 |  |
| Qy | 601   | GACGAGCGGAAACCTCTCTCTATACATCTACACACCGTCGTGTTCTTTTACTGGTTC      | 660   |  |
| Db | 30392 | CACGAGCTGAAACCTCTCTGCTCCCATATTTTGACGGCGGTGGTGTCTTCTATTGGTT     | 30451 |  |
| Qy | 661   | CTAAACAGCTGGCTAAATACCCCTCTTTACGCCAGGGGGGTTTCGTGCTAGATATGC      | 720   |  |
| Db | 30452 | TACAAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTGGCGTGCCTTGATAGTCTG    | 30511 |  |
| Qy | 721   | GTCCGTTCCGCAACGACAGCAAAACGCTTCCAGTCAACTTACAAACACGCGGACGCAAAAG  | 780   |  |
| Db | 30512 | GTGGGCTCTCCACCCAGACGACGCTTTCAGTGGTCTTCAACCAACACGACGACAGA       | 30571 |  |
| Qy | 781   | ACGGAGATATCGTCGTGGAGAACATCTCAATCATATGATAGGGAAGAGGTGCTATCT      | 840   |  |
| Db | 30572 | AACGAGATATCTGCCCTGTGATAATGTGCTCACTTCTCTATTGGGAAACCGTGAAGTTGCA  | 30631 |  |
| Qy | 841   | GAGTACGTACACTTGGAGAGAACTGGCGCTGTTTATAAACAACAGCGCTGCACCGAT      | 900   |  |
| Db | 30632 | GATTATATTAGGCTGGACGAGATTGGGCCCTATTATTAAACAATAGCGCTGTACCGAT     | 30691 |  |
| Qy | 901   | CTGCAGTTCCGACACCTTTCTAAATAATTTTAGAGTATATCTTCCCATCTCTGACGAAATC  | 960   |  |
| Db | 30692 | CCCCAGTTGTGCTACTTCTGAAGACCTTAGAATATATCTAGACATATCACAGAGTTA      | 30751 |  |
| Qy | 961   | ATGAGTTACGTAGACAGATTTGTCTCTCTTAAAGTAGATATTGGATCTCTTTTAGAATAC   | 1020  |  |
| Db | 30752 | ATGGACTATATAGATAGTTTGTGGTTCGAGAGTAGATATCTGACCCGCTCGATAT        | 30811 |  |
| Qy | 1021  | ATTGGCTGACCCGACTCTTTTGTACATAGCGAGGTAAAGCGTATCTTAACAAACCTA      | 1080  |  |
| Db | 30812 | GCAGGTGGACAAAGACTCTTCTATCTCACACAGGAGGTGAAGTCTTTCTGGCAACGCTG    | 30871 |  |
| Qy | 1081  | CACACATGTCATCGCTAGGGGGCGATACCGGGACAGAGCTCTTTACCTGTCCCGTG       | 1140  |  |
| Db | 30872 | CACACTGCTCTCGAGTAAATAGGATGCTGTGTCCAAAGCTTTTCCCTGCCCCAGT        | 30931 |  |
| Qy | 1141  | GTGTGCGAGGTGTTGTGAAGCCGTTTGAAGGAATACAGCGGGCGCTCAACTCACCAC      | 1200  |  |
| Db | 30932 | GTCTGTAGGTGTTTTACAGAGCCATTTTGAGGAGTACAAACGGCGGTAGGCTCACAC      | 30991 |  |
| Qy | 1201  | CTCACCGTGACCGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAATTTTCGAGTTT      | 1260  |  |
| Db | 30992 | ATGACTCCCATAGAAATGGGTAAACAAAAATCTTTTCAGCTAAGTAACTACTTCGAGTT    | 31051 |  |
| Qy | 1261  | GTGGACCAAGGACATGTCATAGTGCACAGGAATCCACTGAACGTTTCGACCCAGGTACC    | 1320  |  |
| Db | 31052 | GCTGATCAGGACATGGCTGTGTGGGACCTATATACAGACGGCTCCACACAGATACC       | 31111 |  |
| Qy | 1321  | TTTATACCAAGTTTGTAAAAACAGCCACGCTGTCTTTAAATGGAAAAACAAAAAATGC     | 1380  |  |
| Db | 31112 | TTCCGCACTAAATTTGTCAAAAACAGCTATGCTACCTTACTGTGAAAGACCAAAAAATGT   | 31171 |  |
| Qy | 1381  | ATATGTGGGTTTCAGGGTACGTACTTTTGTGTTTCAAAAGAAATCTTAGACAGGAACTTTC  | 1440  |  |
| Db | 31172 | ATATCGGGTTTTCAGGGTCAATACCAAAAGATTCAGTCCATCTAGACGGGAGCTATTT     | 31231 |  |
| Qy | 1441  | GTGGAGACGCAATTCGCAAGACCGTCCGGAATAGTATACGGTTTCTTAAACACGCTACTG   | 1500  |  |
| Db | 31232 | ATCGAAAGTCAATTCGACAGATAACCCCGCTTATGTGTACAGTTTCTTAGTACCTGTGA    | 31291 |  |
| Qy | 1501  | TACAAACGCCATGTACTCGTTTTCACGCTACCGCGTACGTAGTGTACACAGAGAAATATCTG | 1560  |  |
| Db | 31292 | TATATGCCATGTACTCATTTTACGCGCACGGGGTGAAGCAGGGGCATGAAGAATTCCTC    | 31351 |  |
| Qy | 1561  | CAAGACCTTAAATTTGCAACCCCTCCGCGCTCTGGCAACCGGCGGTGTAGACCTTCAA     | 1620  |  |
| Db | 31352 | AGGGACCTCAGGAACTCGCGGTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCAGGAC     | 31411 |  |
| Qy | 1621  | ACGGTTCTGTAAGAGTTAAACCTGGAAGACGACATCTTCTACACGTTGTGAGTCCCCCG    | 1680  |  |
| Db | 31412 | GTCTCGGGCAGGAG-----GGGACACAGATGCTTCTACCTCACCGCAGCGCTCCA        | 31465 |  |
| Qy | 1681  | CCACCGCGGGTATCACTCCCTCCAGGTTTGTGTCGACAGTACTGCGCCCTTAAGGAC      | 1740  |  |
| Db | 31466 | CCATCCCGCCACCGCGGCTCTTCCAAACACTGGTGGGCTATTACTCGGGGCGCAAGGAA    | 31525 |  |
| Qy | 1741  | GTGTGCGCTCCAGAAATAAAGGTGGCTGTCTGCTGTTTGGCGGGAGTTTGAGAAGGAA     | 1800  |  |
| Db | 31526 | CTATTCTGCAACAGGCTGGCCCTGACCGCGACACTTTGTGTCAGAGTTCTCTCCACTCC    | 31585 |  |
| Qy | 1801  | AGTTTTTCGCGTTTACGTTTAAACATGTGTCGTAAGGACGAGTTGACTTGTCTCTCCCT    | 1860  |  |
| Db | 31586 | GATTTTTCAACGTTTACGTTGAACATCGTGGTGGAGATGGCGTGGACTTTGTGCTCAT     | 31645 |  |
| Qy | 1861  | TCAGAACGCTCAACGGGCTGTGGCTTTGTCATCGACCGTTTGAATTCGTATATAAATTAAG  | 1920  |  |
| Db | 31646 | TCGCCGGCTCCAGGCTAGTGGCATACGATCCACTATAGACACCTATATATATCCAG       | 31705 |  |
| Qy | 1921  | GGGTACACGTTTTTACCGGTACGTTTGGTCTGCTGTGAGGTTTGGCTCTCATGTATGAC    | 1980  |  |
| Db | 31706 | GGATATACGTTCTCTCCAGTGAGATTCCGCGCTCCAGGAGGACAGCGCTCAGCGAGGAC    | 31765 |  |
| Qy | 1981  | CTCAGAAAGATGCCCTCCCTGCTGTCGAGGACTCTAGCGGTTTATTCGCGTGCCTTA      | 2040  |  |
| Db | 31766 | CTGGCAGAAAGATGCCCTCCCATAGTTTTCGAGGACTCATCGGGTTTCAATTGCTGCTG    | 31825 |  |
| Qy | 2041  | GAGAATAACATAACCAATTTGACCGAAACCATGAGGAGCGGAGCGTTTTTCAAGTGTGC    | 2100  |  |
| Db | 31826 | GAAAATAACGTCACCAAGATGACAGAGACCTCGAAGGTGGCGACGCTGTTTAAACATATGT  | 31885 |  |
| Qy | 2101  | TGTGCGGGGACTTATGGGTCAGCTCAAAATTTAGCCATGACCATCTTAAGGCACAGGGA    | 2160  |  |
| Db | 31886 | TGTGAGGGGACTTACGGTATCAGTTCTTAATCTGGCTATGACCATAGTAGGACACAGGG    | 31945 |  |
| Qy | 2161  | ATGTCGTTGGAGCGAGTAGCCGTAGTATTTGGGTCCCAAGAAACGTCGACAGACCCAC     | 2220  |  |
| Db | 31946 | GTTTTCAATAGTAGGTGGCCATATCGTTTCGGCAACACCGCAATATCAGAGCCAGTCTA    | 32005 |  |
| Qy | 2221  | GTGTATGTAGCAATATCAAGGCTGTAACTCAAAATTTTGTGTCTATGACAGCAACCC      | 2280  |  |
| Db | 32006 | GTGTATGTGGTGTATCCAGGGCCCATCGACGCTCTGTTACTGTATATGACAGTAACTCC    | 32065 |  |





RESULT 10

EHVU20824 184427 bp DNA linear VRL 02-FEB-1996

Equine herpesvirus 2, complete genome.

U20824

U20824.1 GI:695172

Equine herpesvirus 2

Equine herpesvirus 2

Viruses; gDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae.

1 (bases 1 to 184427)

Telford, E.A., Watson, M.S., Aird, H.C., Perry, J. and Davison, A.J.

TITLE The DNA sequence of equine herpesvirus 2

J. Mol. Biol. 249 (3), 520-528 (1995)

MEDLINE 95302501

PubMed 7783207

REFERENCE 2 (bases 1 to 184427)

Telford, E.A.R.

Direct Submission

Submitted (09-FEB-1995) Elizabeth A.R. Telford, Institute of Virology, MRC Virology Unit, University of Glasgow, Church Street, Glasgow G11 5UR, United Kingdom

LOCATION/QUALIFIERS

1. 184427

/organism="Equine herpesvirus 2"

/mol\_type="genomic DNA"

/strain="86/67"

/db\_xref="taxon:133893"

1. 17553

/db\_type="direct"

9248..10399

/note="ORF E1"

/codon\_start=1

/product="G protein-coupled receptor"

/protein\_id="AAC13788.1"

/db\_xref="GI:695173"

/translation="MATTSTATSTNTSLATTMTNFTSLTTSVVTIASLVPSTNSN EYDDDDVDVEEAPCVTSQVTLAAQVLPALYVLFGLGLNLIIVITIRYMK IKNTLMNLNLAISDLLELLLPFWHVIYGMHWDTFGISLCKLRGVCYMSLYSQV FCILLTVDRYLVAVYAVTAFTVTCGIVCTVCTAGLISLPEPFEGHQDND RVQCPDYEMSTNWRRAHAKVIMLSILLIMAVCYVILIRLRSPSKKYKA IKLIFVAVGAEKFRRLHFFHTYVAILYCKYIPFLSGGEGKRGFTRI"

10494..10499

/note="ORF E1"

/note="ORF E1"

10660..10665

/note="ORF E1"

12533..13114

14197..14541

18810..19733

/note="ORF E2; product has hydrophobic N-terminus"

/codon\_start=1

/protein\_id="AAC13789.1"

/db\_xref="GI:695174"

/translation="MIILGLVLSLASTASAIQAWMPQVITGHERGPLSVRCHVAGK NSPKPMVLLKTIGSAPRGITVPFKGVTIKSSCSGTGVTFTVDSLHMADAGRYA CWSQNDPKDSVCSDDYATNTIVLETRRRSLIGQLPTAGPSTOLSTQTP LSSQESFVAVRGKVIYGVSLVEVCEQFSAGTGLNISAVLPKKVGEAGYQAVST ISTSLTLDLSVIEGTQAKWCNVATKDSGEYMCIMGDMDFPEGESEGDDEDFSG VFFKTVTPEGEDADEDDYLYTSEEEKK"

19761..19766

/note="ORF E2"

19784..19789

/note="ORF E2"

20320..20844

/note="ORF E3; product has N- and C- hydrophobic domains and 6 N-linked glycosylation sites"

/codon\_start=1

/protein\_id="AAC13790.1"

/db\_xref="GI:695175"

/translation="MATEVGFACALVLLGCGVAGTPTPSKTLIYNSONCTTYSIE NGOSLTVNGSWPIRVEFNCSSGVELOGWPTTCTIENKNGITWNGPSSCVKLNITV

LCSSLLPCLRLQNDTSCPLSLLELOHLQNLONTREALGVLTDLLSGDPSRSEYFEALHPS  
 RPLGCKRHARVFPFYGAETKSLSLNDVEFFFKELNSVFCYCLPAEGALEALGE  
 TVAFGLRGVSLIPADAYSSVPCASGCFSAAMLPNQGESVLSMLAAVNCVNCRCQ  
 VPSDPVIGVFNELHLAGADARAAGCERRERREDADDEDEDEDEDEDEDEDEDE  
 GMAALRVLTESLSLWAGTIPEDDGRABISNLVYSSAARDGGVGGTARATSSS  
 HMAKFAHARWERSAWLGRGAPSEFPDRAIPSPSLFCG3VFNISDDTAAALQKD  
 CSATFLKSNYQTLIOQNELVYRLEVLNAGRGDEGAKGAEAAADAEPLKPDGAS  
 CDRPDLSDARVARDLYLKLITRDGLRLITDCIETHGRVLSDLSLWVMSYASAA  
 RLNVHLEFRFRQVGLGADLTAGEAAAFENSKITNALHQRLNREHLDLSIVHYRRL  
 ITGPLSLQNSHPFVPPNVALAYCLDAAGAPHQKLVITEMIWPIESKDWIDCNFNSF  
 YSTETGLMLTKKTNYIREALVSLSLYVNRKSLSLSLATSLRGLSCLABSSGEL  
 GEGVLYLTGEGTAPLVLFVDSKGVFKDLVTLVLYTHQLSGRRQASV"  
 30114..30119  
 /note="ORF E3, 03, E4 and 06"  
 /notes=".34819  
 /codon start=1  
 /product="glycoprotein B"  
 /protein\_id="AAC13795.1"  
 /db\_xref="GI:695180"  
 /translation="MSVGGPRVVLCLWCAALLCGVAQEVVAETTTTPFATHRPEV  
 AENPANPLFPFVFCASPTGEGEIFFPLESCPNEDKHIEGLAIKTNIVPIFV  
 NVKRYKIMTSTIYKGBEDATNQHTRSYAPLYEVQMDHYOCPSAVQNEGGH  
 VNTYDRDQWNETAFILKPADGLTSSITRYSQSEVATPRNLLWSYTRTTVNCVTE  
 MSARMKPEFFVTSVGDTEMSPLFKNGTSEPEKILRPHSILQLLNKVAVTGCVGL  
 GQADNATRPFAIPGDYSLWSKATTENSSYCDLILKGFNSAIQTOHNSLHPLANDIT  
 ASPTPLEEENENETPKCIWNNTORSEIOKLLKEVEKTRPNGTAKVYKVTGNIYLIV  
 QPLQIDLDTHAKLINLNAVASPTSTETTPRRRRDTSVSGGNGDSTKES  
 VASQVQFAYDNRKDSINVLGELSRWCREQTRASLMWELYSKINPTSVMIAIYGR  
 VSAKLIGVDSVDCISVQSVFVKNNKFKEDLCYTRFVGFNGFNGSELFAQ  
 LSPRNIIVLSTQVEVQSCHEHYFOAGNQMYKDYVYVSTLNTLDITPLHTMITLN  
 LSVENIDPAVILEYKTEKRLSNVDIETPREYNYTONLNLRLKDLIDHGRD  
 SFTQITGLDMQDGLTGKVVANVASGVPSLFGISVSGVLSFFKPNFGMLLVLIAG  
 VVVYLPMTRESRIYAPILMFLYPGVVERAAQEPGAPVSEDIQRLNLLMGHPQORQ  
 AEEBAREEVRKRLTFEVRDTSATSVLRRRGGGGYQRLQDGDGDEDEYEPURQ  
 DGGYDDVDVEAGTADTV"  
 34855..34860  
 /note="ORF 07 and 08"  
 /notes=".38115  
 /codon start=1  
 /product="DNA polymerase replicative subunit"  
 /protein\_id="AAC13796.1"  
 /db\_xref="GI:695181"  
 /translation="MSGFNPYLVRKTLFLKKAAPSRPTKEYTRIIPKCFKTPGAAGVVP  
 HTSTLDPCVPGNGPILYDGSRLWSAGRGPGTGAGGQHTPVALTFFHYDIE  
 LDTYQDCRHVPQFOTDIIIPSGTVLKLGRSDDRSVNVFROELPYFVRPGLK  
 LPLVQCSENEPNSQGRYVEKTSKVLRYCVBARVBYVFASSGQFVLLIAGEL  
 TAAGCEFEINVDARFIIDNGFTFGYSCAAVPROGGAARSWTELYDCAAD  
 LEFHAGRMFGNVLVSFDIEGNGFPNABRDEMIQISCVIWKGASGAPRSLV  
 LNLGTCEELEGVEVYQPSBELDLYLFTMRDADVEFTGYNINFDFFYVIDRATQ  
 VYNLKTEFVRASSSIFEVHKPNSSAGFMRAVSKVAVGVVPMIDMYQVCRDKLSL  
 NYKLDTVAGCEKEDVSYKEIPLHFRQPGGRARILGLYCVKDSALVLDLIRFMT  
 HVRISEIAKTIPTRRVLTGQOIRFVSCLLDVAGREGVILPVRHADAEYQCATV  
 IDPSRPFYNTPLVDPASLYPTIIQAHNLCSYTMIPGDELCLHPLHGPQDYETELA  
 SGPVHFKGHKVASLLATLLNWLAKRKAIIRLRELTASDEAVRTIILDKQQLAKVTGN

Query Match 10.7%; Score 974.6; DB 14; Length 184427;  
 Best Local Similarity 54.2%; Pred. No. 5.2e-246;  
 Matches 2289; Conservative 0; Mismatches 1859; Indels 79; Gaps 12;  
 34 GAACCAATCGGGGTTTCATCTTGACATGACCTCGATGCCAAGTCAGGGGTGCTG 93  
 92963 GAGCCGGGGCGGGGTTTCATCTTGACATGACCTCGATGCCAAGTCAGGGGTGCTG 93022  
 94 GATCACGTGAGTGGCTGTCAATATATACTACAGCCACCGGAATGGGTGGTACGAC 153  
 93023 GGCAGGATCGGGGCTGGCGCAAGACCGTCCCGCCCGCCGAGATGAGTGGTTCGAC 93082  
 154 CTGCGCTTCGATCGGCTGAAGACTCGGCGCGCTTCTTGGCGCTTACCGTTTATCTAAT 213  
 93083 TCTCAGTTCGACCCCTGAGGACGGGGGCGCCCTTCTCCCTTCTCCACCTTCTCTGATA 93142

214 ACGGAACTGCTGTGTGGGAAAAGTACACAGCATATCGGCGCTGTACCAAAATTTAAAC 273  
 93143 ACGGCAACGGGGGGGGGTAAGAGCACTAGTATCTCGGGCGCTGTATCAGAGCCTCAAC 93202  
 274 TGCCTGATACAGCGCGGACCAACCATAGCCACACAGAACCTATCGCGCTCGCTTAAAGACG 333  
 93203 TGCCTCATCACCGGGGCCACGGCCGTGGCGCTCAGAACCTGTCCAAGCGCTCAAGACC 93262  
 334 TTCTGTCCACGATCTTCAGCGCTTTTGGCTTTTAAAGCGCAGACATCAATATATAGCGTC 393  
 93263 TACTGCCCCACCATCTACAGCGCGTTCGGGTTTAAAGACGAGCACTTAATATATATGCGC 93322  
 394 AGAAAGCTCATCAGACCGGAGCGTATCTCATAGACAAATTCAGCAACAGGAGCTATCG 453  
 93323 CGCCACGGCCGAACGGCGCGCGGACATGAGGGGATCCAAGGAATGAGCTGTGC 93382  
 454 AAGTATTGGCGGTTATAGTGACATTAATAAGAGGTTATGGCGAAACCAACCAATGGC 513  
 93383 AATACTGGCCGTGATCTCAGATATATCTCGGGGAGTTTACCAAAAAGACAGCGGGGC 93442  
 514 ATGTACGGGACTATATCCAACGGCAATTTTGAACCCCTCTCGAGNATGACCGGCGGT 573  
 93443 CAGTACGAGCACCTGACGGGGCGCGCTTTGGGGCGCTGGCAGATGGGCACCCCAAC 93502  
 574 TTATGGACTTCCAATATATTATGTAATCGACGAGCCCGAAACCTGTCTCTTATCATATC 633  
 93503 CTGTGGACCAACAATCATTTGTATAGACGAGGGGTACCTGTCTCGACATCTGT 93562  
 634 ACCACGCTGCTTTTACTGTTTCTTAAACAGCTGGCTTAATACCCCTCTTTACCGC 693  
 93563 ACGCGGTGTATCTCTGTTTATCAACAGCTGGCTGCGAGACCCCTCTGTACAG 93622  
 694 CAGGGGGGGTTCGTGTCATAGTATGCTCGTTCGCAACCGCAGACAAAACGCTTCCAG 753  
 93623 AGCGGGCGGTGCTCGGTGCTGCTGGGTGCTGCGGCGGACGAGCGGACCATACAG 93682  
 754 TCAACTTACACCGGAGCAACAAAGACGAGATATGCTGCTGGGAGAAATCTCTTAAACA 813  
 93683 TCCACTTACAACTCTCCATGCAGAAACCAACATACAGAGTGCAGCAACATCTCTACC 93742  
 814 TTCTATAGTAAAGAGGCTGATCTGAGTACCTGACATTTGGAGAGAACTGGCGCTG 873  
 93743 TTTCTCATGAAGACGAGGCGGTGTCCAGGTACGTGGACCTCAACCACTTGGCGCTG 93802  
 874 TTTATAACAACAAGCGCTGCACCGATGTCAGTTTCGAGTTCGAGACCTTCTTAAATAATTTAGAG 933  
 93803 TTTATCAACAACAAGCGGTGCACGAGCCAGAGTTCGGGCACTGTCTCAAGACCTCTCAG 93862  
 934 TATATCTTCCATCTCTGACGAAAGTATGAGTTACGTAGACAGATTTGCTGCTCTTAA 993  
 93863 TACAACTTGACATCTCCCCCGTATGAGTACATCGACCGCTTCTGCTGCTGCCAAG 93922  
 994 AGTAAGATTATGGATCTCTTTAGAAATACATTTGGCTGGACCCGACTCTTTTGTTCATAGC 1053  
 93923 AGCAAGATCTGTCCCCCTCGAGTACGTGGGCTGGACCGGCTCTTCTGCTGCCACAGG 93982  
 1054 GAGTAAAGCGGTATCTTAAACAACTTACACATGTCTAACGTTAGGGGGCGATACAGG 1113  
 93983 GAGTCAAGGCTTACCTGACGCGCTGTGCAGAGACCTGTCCCTGAAACCAAGGTTGGCGC 94042  
 1114 -----GACACGAAGCTCTTTTACCTGTCCCGTGGTGTGCGAGGTGTTTGTGAAG 1161  
 94043 ACCCGGAGGCGGACGCGAGGCTCTTACCTGTCCCGTGGTGTGCGAGGTGTTTCCCGAT 94102  
 1162 CGTTTGAAGAAACAAGCGGCGCTCAACTCACCAACCTTACCGTACCGAGTGGGTG 1221  
 94103 ACCTTCAACGAGTACAGGAGCGGTGAACCTCGCGGGGCTCAGGTCACCGAGTGGGTG 94162  
 1222 ACNAAAACCTTTTAAAGTTAAATTTATTCGAGTTTGTGGACCGAGCATGTCCATA 1281  
 94163 CAGAAGAACCTTCTGAGGCTCAGCACTACTCCAGTTTATAGACCGAGCATCTTCCGCC 94222  
 1282 GTCCGCCAGGAATCCACTGAACGTTTCGACCGAGGTTTACCTTTATCAACAAGTTTGTATAA 1341

polyA\_signal

CDS

polyA\_signal

CDS

Db 94223 GTGCACATAGAGACGCGGAGGATCCACCAAGGTCACTTACCTGGTCAAGTACGTGAAG 94282  
Qy 1342 AACAGCACGTCGTCCTTTAAATGGAACAAACAAATGATGATGCGGTTTCAGGTAGC 1401  
Db 94283 AACAGCTACGTCGTCCTTGAACGCGCAACCAAAAGTGCACTCTCGCGGTTTCATGGGCACC 94342  
Qy 1402 TACTTTGAGTTCAAAAGAAATCCTAGACAGCGAACTCTTCGTGAGAGCGCATTCGCAAGAC 1461  
Db 94343 TTTGAAAAGTTTAAGACCATCCTTAGACAAAGAGACCTTCATAGACGCCCACTCCACGAC 94402  
Qy 1462 CGTCCGAAATACGATATACGTTCTTAAACACGCTACTGTATACAAAGCCATGATCTCGTTT 1521  
Db 94403 CAGCCCGAGTACGTTGATACGTTCTTCAACACCTCTGTATACAAAGCCATGATACGCTTC 94462  
Qy 1522 CACGCTACGCGGTGACTAGTACACAGGAAATATCTGCAAGACCTTAAATTTGACACC 1581  
Db 94463 CACAAGCAGCGGTGAGACGCGGGGACGAGGCTTCTGACGCGCTGAGGCGGTGCGCC 94522  
Qy 1582 CTCGCGCGCTCTGGCAACCGGGCGTGTAGACCT-----TCAAACGGTTCGT 1629  
Db 94523 ATACCCACACATACCCACCTCAGCACGTTCCAGGACGCCCTCGACACGACCGAGGCG 94582  
Qy 1630 GAAGAGTTAACTGGAGAGGACATCTTACCAAGTGTGAGTCCCCCGCCACCGCG 1689  
Db 94583 GCGCTCCTCAACCCCGAGAGGACATCTTCTATACATGACCTGCGCGCCCCCTCCGCC 94642  
Qy 1690 GGTATACCTCCCTCCAGGTTTGGTCGACACGTAATCGCCCTTAAAGGACGTTGTCGCC 1749  
Db 94643 TCCTCGCGCTCCCTGTCCACCTCATCTCTTTTACATGTGCTCAAGTCGCTGTTCCCTG 94702  
Qy 1750 TCCAGAAATAAGGTGGCGTGTGCTGGTTGGCGGGAGTTTGAGAGGAAAGCTTTTCC 1809  
Db 94703 CAAAGGTCGCGCTCGGGGCTCTCTCGCTTCGGGGGAGCTTTCGCGAGCGACCTTTTCAG 94762  
Qy 1810 GCGTTTACGTTTAAATGTCGTAAGGACGAGGTTGACTTGTCTCCCTTCAGAACGT 1869  
Db 94763 ACCTTCAATAAATGATGATACAAACGCGGTGACCTTCACTTCGCTCCGAGCGG 94822  
Qy 1870 CTCACCGGCTGTTGGCGTTTGCATCGACCGTTGAATCGTATATAAATAAGGGGTACAG 1929  
Db 94823 CTCCTTGGGCTCCTGGGCTACGCTCCAAAGTGGACACCTACAGCTCAAGGGCTACACC 94882  
Qy 1930 TTTTACCGGTAGCGTTCGCTGCTGACGCTGTCAGGTTTGCCTCACTGATGACCTCAGGAG 1989  
Db 94883 TTCATACCGTGGGCTTCGCGCGGTTCAACGAGCCGAGCTCAGCGGACCTTCGCGCAC 94942  
Qy 1990 AAGATGCCCTCCCTGGTGTGACGACTCTAGCGGTTTATCGCGTGCCTTAGAGATAAC 2049  
Db 94943 AAGATGCCGTTGGTGTGAGGACCCCAAGGTTTATAGCTTCATAGCTTCGCTGAGAACAC 95002  
Qy 2050 ATAACCAAATTGACGGAACCATGAGGACGAGCGGCTTTTCCAGTGTGCTGTGCGGG 2109  
Db 95003 GTCACAAAGATGACGGAGGTCTGGAACACGGACCTGATACACATCTGCACGGCGGG 95062  
Qy 2110 GACTATGGGTCAGCTCAAAATTTAGCCATGACCATCGTAAAGGCACAGGGAATGTGTTG 2169  
Db 95063 GACTACGGGATAAGCTCCAAGCTGGCCATGACCATAGCCAAAGGCCAAGGCAATGTCCCTG 95122  
Qy 2170 GAGCGAGTAGCGGTAGTATTTGGGTCCCAAGAACGTCGACGACCAAGCCAGCTGTATGTA 2229  
Db 95123 AGCCGGGTGGCGCTGCTGCTTTGGCAACAGCAAGTTTGTGAGAAAGACCCACGTTAGCTG 95182  
Qy 2230 GCAATATCAAGGGTGTAACTCAAAATTTTGGTTCATGGACGACCAACCCCTTAAAC 2289  
Db 95183 GCCATCTCAGGGCCACAGACCCAGACACATGCTCATAGCTGTAAACCCCTCAAGAAC 95242  
Qy 2290 CTCCTCAGAGAACGATCGATAACACTTCGCGCCGAGCATATAGTCCCGCGCTTCACAAAC 2349  
Db 95243 CTGAAAGAGGACCGCGAGGACAAACAGACATCCCAAGTACATAGTTCAGCGCTCAACAC 95302  
Qy 2350 CCAACACACCCCTCATCTACTTAATAATAAACAAGCAATGGAATAACATTTGTTT 2409

Db 95303 CCAGACACTACCTTTGTTTACTGAATGGAGATA-----CCACTACAAAAAATTTTACAG 95356  
Qy 2410 TTATTCAGTCCAAACACGCGCACCGATAGTTGTTCATTTCCACACACCGGGGGTGGCG 2469  
Db 95357 ATTATCTGTTATGTTTCCCAATATAAAACATTAACCTCAAAATCTCGTTTAATTTTT 95416  
Qy 2470 CCATAGTTTACGACACGAAACGCGCGGCTCTGCGTTTGAAGTAGACCCGCCGCCCG 2529  
Db 95417 TTTATTTTATTAATTAATCAATGGGAGAGAGATAGAGAAAGAAAAGCCCTTAG 95476  
Qy 2530 TGGTCAGGTCATACGTCCTTAGGGCCCTGGTTGACCAACATAAATACGTCGTCAT 2589  
Db 95477 AGTTTTCAGCAAGGCCAGTT--GTAGGTGGGTTCCCGGCCACAGGGGCGTCTGCGC 95534  
Qy 2590 CACCGTTTCAGTCCCTGCGTACGCGCGCTCTTTCAGGCGGCGACGAAACACCGCGCTG 2649  
Db 95535 CA----TATAGTTTTCGATGACTCTCTGGACATGACCAACAGGACCTGATCTCTCGGCC 95590  
Qy 2650 TAACTGATCTGCATCTGCTGCTGCTGCTGAAGATGAACACAGTCAATCAATCAGTAGG 2709  
Db 95591 CAAAGAGTCCCTCCCTCCCGCGCGCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 95650  
Qy 2710 CGCCGTCCTCATCGCGCTCTGATTCGTCCTCCATGCTCAACAGAGCTCTCGTCTCCA 2769  
Db 95651 AGATATAGATCACTTGTGTTGGGGTC-----ATTGCTCGAGCTCTGGAGTGT 95704  
Qy 2770 CTTCGTCCTATCTGAATAATAATGTTCTTGTCTGCTGAGTTGATCCAGCAACACAG 2829  
Db 95705 GCTTCTCTTTTGAAGCGCGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 95764  
Qy 2830 GTTCTCTTTCATCGAAGACCATCGCCCTCTGCGGCTGATCTAGATATATACCTGACC 2889  
Db 95765 TGTCTGACTCTCTCGAGG-----GAGCAGGCTGATGACAGGACTTCCCTTTGCGC 95817  
Qy 2890 AGCGCGGCGAGCGCGTTCGCGGAAAGCGCTCAGCTTCTTCGCGCACTTTTCATCAGTGTCT 2949  
Db 95818 GGAGGGGTTTGGCGCGGGGGGAAACTCCCGCGCGCAGCGAGCTGCTGACCTGCGCTC 95877  
Qy 2950 CGTACTCATCTATAAATAAATGCTCTGCTGCTGATTCAGATGACTGCGGTGCGCGGT 3009  
Db 95878 AGGATATCTGGCTCGGAGTGTGCGATCGCCATCGCTCTCGTGTCTGCTGCTGCTG 95937  
Qy 3010 GTTGGCCAGTAAACACCGCGCGCGCGCACCAACCAACCGCGCTAGTGGCGCGCGCG 3069  
Db 95938 CTTTCGCTATCCGAAGAAG 95995  
Qy 3070 CGCTCGGTGTTGTTGTCGACCTGTGCAAGTAATGACGCTAGGTAGGTAGCGCGCTCGT 3129  
Db 95996 AGCCCGGCTATCTTTATCCCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 96055  
Qy 3130 TAAAGGACTCACCGTCCGCTGCTGCGGATCTATTCTAGGTGGGAGAGACCGCGGTT 3189  
Db 96056 TGTCTGCTAGAGCCCCCGGGGAGTCTCTGGAGTGCAGCGTCTCTCATCGTCTCT 96115  
Qy 3190 CGCTCTTTCATCCATCATCTGATGTCGTTCTCCGAATGAGAGTGGTTCGTTCTGCTG 3249  
Db 96116 CCCCCCGTCTCTCGGAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 96175  
Qy 3250 GGGTTTTGGAGTAGTGGAGGGGAATCTTGTCTCAAAACATCTCTCAGGTGCAGTGT 3309  
Db 96176 GGGTTTTGGCTTACTTAAACTGGAACGTGGGAGGGGGTTGCACTGGGTGCAGTCT 96235  
Qy 3310 CAATCATCTGGGGGCTCTCGTACCCCGCGGTACCGGCAACCCCGTACGGGAGCC- 3364  
Db 96236 CAAAAAGACATCATCCAGGTCCCGCGCTTTTAGTCTTCTCTCATAGTTTCTTAATCC 96295  
Qy 3365 -----GSCGAACAAAAACGCTTGTCTCTTTTCTTAGGTGCCCGGGAATCGGCA 3417  
Db 96296 CATCTTTGGGAACCTGAAAGTAGTGGCCCGCTTTAGCTGGCGCCCCCGGTGCTG 96355  
Qy 3418 GCATCTCGGAGTGCC---GSCGGGTCTTTGAGCGGTGCGATGAAGACATAGCAATGG 3474  
Db 96356 TCATCTGGGTGGCGACCGGTCGCGCGCGCGCTTCTCTCTCGAGGACCCCGCTTAG 96415

|            |  |   |                            |
|------------|--|---|----------------------------|
| QY         | 3475   | CGAAACGTTACCTGTAAAAACGCACTGCCTAT-----CCCCGATCAGTTAATAT        | 3524                       |
| Db         | 96416  | TCCTTTCTGCTTTTTCAGAAACATAGACATGTCGAGAAAGCCCCAGTCAGTCACAGGT    | 96475                      |
| QY         | 3525   | TCCAGTCAACGGCACCCCTCCGATTCTGAATAGATAGTCAATTCGCAAGTTTAAAAATGCT | 3584                       |
| Db         | 96476  | TCCAGTCCACCGGACACATTCGCGTGCACTCCAGTACTCGTTGCGCTGTGTTAAAGTGAC  | 96535                      |
| QY         | 3585   | TGCAAGCTAGAACTGCGGCCACGGAGAACCGGTTGCGCGTGGCGCGTACGCCAGCGGAG   | 3644                       |
| Db         | 96536  | CGACGCCCAAAAACTTTGSCACAGGGAA--GGTCTGTCTCTGCGACGCCAGAGGGG      | 96592                      |
| QY         | 3645   | ACGGATGCGCGCTTGAGGACGAGATCTCTCTGTCGTTTATCATCAGCGTGCCTTGTCTA   | 3704                       |
| Db         | 96593  | ACGGGTGCTGGCTCTTCAACACCATGCTTGTAGTATTTATAGACTGGCTTTTCTA       | 96652                      |
| QY         | 3705   | TAGCCCTTTCTTCCCCACAGCATAAACGCGACCTCTGAAGCTTGCGCAAGAGCGAGCTTA  | 3764                       |
| Db         | 96653  | TGCGCTTTGACCCCAAGCATACACACAAATGGTCAAGCTCATTTTGAGAGGCGAGCTGA   | 96712                      |
| QY         | 3765   | TTATGTAGTGTGTGAACCAATCCAGCGCAAGGTTGGGTGTGTGACCCCGCTTCCCTCTTT  | 3824                       |
| Db         | 96713  | TAATATAGTGTGTGAACCATGATCCACCCCAAGTCACTGTGGAGCCCGGCTTCCCTTTT   | 96772                      |
| QY         | 3825   | CCACCGTCAGATGCTGTTTAAAGCAGAACTCCCGGTTTGGCCCAACAGTCTAAGCAAC    | 3884                       |
| Db         | 96773  | CGACCGTCAGTATAGTGTTAATATAAATCTCCCTGTCTGCCAGTTGTTTAGGCAAC      | 96832                      |
| QY         | 3885   | CGTGAGAGGACGGGTGAACCGAGTACAGTATGGCTATCTCTTTAAAAATATTTTGA      | 3944                       |
| Db         | 96833  | CGTGAAGAGGAACTGCAAACTGTGAGTCACAGCGCAGAGACTTCCAGAAAAATGTTCTCA  | 96892                      |
| QY         | 3945   | GCGTTGGAGGTACTGTAGTCCGGAGCAACGCTAAGGCTAGTCCGCTGGCTTGACCGC     | 4004                       |
| Db         | 96893  | GACTGGAGGGACCGGGTCAGTCTTTTGACACATAAAGCCAGTCCCGTGGCTGTCCCG     | 96952                      |
| QY         | 4005   | GATGTGAAGGATCTCTGGCCAGGATCAAGACTTTAATATCTCTCCGGTTGCGCAACATACG | 4064                       |
| Db         | 96953  | CGTGATAGGGTCTCTGCCCAAGATTACCTTAACCTGGGTGGGCTCGCACAGCCAGG      | 97012                      |
| QY         | 4065   | ACGACACATAATCTATCTATCGTGGTGAATAATCAGGTTACGTTGCTCATATCCATAA    | 4124                       |
| Db         | 97013  | ACCAAGAAAGACCCCTCTCTCCCGGGTAGACACCGTAGAGTCTCTAGTTTGGCCA       | 97072                      |
| QY         | 4125   | CGCGCTTCAGCAGACGGCGAGCTTCTGTTTTTAAAAATGGAGAAAGATTAAAAATCAA    | 4184                       |
| Db         | 97073  | CAGTCTCCAGCAGAGTCACGACTTCTCTTTAGAAAGGTGACATCTCGAGAAAGTCCA     | 97132                      |
| QY         | 4185   | GCCATGAGTCGTAAAGTAGCAAGTTT                                    | 4211                       |
| Db         | 97133  | TCCAGGCTCTGTTAATTGGAATCTTTT                                   | 97159                      |
| RESULT 11  |  |   |                            |
| LOCUS      | AF318573   | 108873 bp   | DNA linear VPL 18-MAY-2001 |
| DEFINITION | Bovine herpesvirus 4 long unique region, complete sequence.          |   |                            |
| ACCESSION  | AF318573 AF271211  |   |                            |
| VERSION    | AF318573.1 GI:12802528   |   |                            |
| KEYWORDS   | Bovine herpesvirus 4   |   |                            |
| SOURCE     | Bovine herpesvirus 4   |   |                            |
| ORGANISM   | Viruses; dsDNA viruses, no RNA stage; Herpesviridae;                 |   |                            |
| REFERENCE  | Gammaherpesvirinae.  |   |                            |
| AUTHORS    | 1 (bases 1 to 108873)  |   |                            |
| TITLE      | Zimmermann, W., Broll, H., Ehlers, B., Buhk, H.J., Rosenthal, A. and |   |                            |
| JOURNAL    | Genome sequence of bovine herpesvirus 4, a bovine Rhadinovirus, and  |   |                            |
| MEDLINE    | Identification of an origin of DNA replication                       |   |                            |
| PUBMED     | J. Virol. 75 (3), 1186-1194 (2001)                                   |   |                            |
|            | 20583805   |   |                            |
|            | 11152491   |   |                            |

```

/protein_id="AAK07925.1"
/db_xref="GI:12802533"
/translation="MAMKTAQGGQVEEDNIGSKAPGPGQYIVYKKEFTPTKEASLIG
NGPFTSVMSVPLLEGLTVEPCGPINVAHKHIDITTVSVKVSREHEVIMFNIDC
FTPIHGTLGALCETRDLFGTHSPSNTNINSIQDLGTCNIDENILGWIT
EGFKSLYHGHMIPLIGHVEVQISFQAKRKIPIDYDLSFKAPDMKRFHSVDVS
QYLFNSLTAQAARAKDAVAVIQAESQFVRDQKMPKVVKKAPDFSSRSGDGL
SLIMDSVLAASVGLSFDIPADQMSALLDYSWPIFVDCITEEDLQALHAWLK
QIHTNLTQFSTNSMLYLRIOKQNPKTAKGLNYSYIQLHGLSYLSEATQBYGQ
VFGQVQSLGSGSTIHLAYAAAMCPNLTARYCYLQFCQHQKSTQNSYINISY
PVQASSDMCNLCQKGPACNTLFLYRDRFPPIITSHRRDPYIITGSGVMYNDLD
ILGNFASPREDEGAQVEVEQVYTWOLTOLEKLGSMGIKONVFAQPCPVFLQ
IYKFDIDAHDEVILKFNCKVKNINLAFRETIKSIHVIQYCONVFAQPCPVFLQ
AEKVMHSDFFCDFDVAAGIIFSPKTIQKISKAMLVVPTKIKIKRIFSSNG
SEAIQGFLLKSGSKDNVITVGYPMKFNINXIMFPQTKMSLFWHTFTSKQIPL
ITGPDQDLVALANYIENSKLHSEIDVLDIDPDLITVAKIKLNNALIRACQTFY
STTLHCVLTFTVPGEEYPHVLENEEFLGTQYLKLVQNRTAQVQATLKEDVAMQ
KRPITVPMVINKYTGINGNGIIFHCNGLYFMGRGVDRNLIPENAPFKQSNTAM
RKGHPMTIVDNLIKRTVTPSSQFEVENIERDMLLLEDKDNONIREDVLELVKG
LGQACADLADDDLOFLYGYIMSBEILSRLOISIDAGVPWCEDSVSYFLGVTVEE
QEFIFLGESQTSAPTSSEDFQASGLSTIAAGKKKLNKMLSDFDL"
CDS
352. .11430
/notes="ORF 7"
/codon_start=1
/product="transport protein"
/protein_id="AAK07926.1"
/db_xref="GI:12802534"
/translation="MARELAAYVSQVDFSLFSTLLTFCDPRHIDLGAFQNKKELKL
LIDALYSHIVEQNETSCSLSELEHLIINTYTWOKLETCLNENIYKQVFMALHS
LDCRFKTNVLCFGVCVPIPSIGSIINDIEIFSLNSVFCISRNALGLEETINF
LGTNRISPIPLDPLVPSISCMNCLNETSILPNOGENVFTMNSINTCHVATPAPE
VOGLPENELHNLINVDGTDNKOVLQETSDTRIQASIEADHNIPEKVSQYLEIS
NLIYNSGHSNDOSPGLSKSQVATLIQHSTRWASHQKVAASHALTPPIHFEDSVRA
YPIESLFGIFYSIMDNIOALQDCSTTFILAKSNKTIEMERKNELYRLNLLAPSH
TKEGDKNKTSIKQARTHITGETTSEQIFSDAAAKNNYLOKVITTEGLIKDCLT
QKQIMQNTLRTWMTYKERAALKNHFMIRKFIALPOWHNTYINALEYSESKFI
KNLSYVQKLSOEHDISITLQFSLITGLVPGDFPSPNILLANCLDAASVMPHHK
MLSEIYVMSPEKMDIDAEAFYNIITGKLSVOCQAMRVIRELWMSVLYNWE
KDLNIVSPSTLSHNLFSKPTFKSGIYIYETKAPLIIHIDNTGWIKDLVALLYH
LQITSKHNVL"
11420. .14044
/gene="gb"
11420. .14044
/gene="gb"
/notes="ORF 8"
/codon_start=1
/product="glycoprotein B"
/protein_id="AAK07927.1"
/db_xref="GI:12802535"
/translation="MYKTIILFALIKVCSFNQTTHTSTTSPSSITSTSTTSK
PSNTTNSLSAASPONTSTKSTDNQGTSTPTIPTVDDTASKNFYKRYVCSASS
SGELFRDLQTCPTDKDKHVEGILLVLKNIVPIFKVKRYKIATSVTVYRGWSQ
AAYNRDDISRAIPYNELSMIDRTYHCFSAVATVINGILNTYIDRSENKSVLPQVA
GLTENIRYSQBLIYABPGWFGIYVRVTNCEVVDVMDVYASVEPYTHFIALGDTI
EISFFCHNNSQCTGNSDRTAKWIEENHQVDYERRGHPTDKRIFELKBEVTIS
WKABDERAICDVIWKTFFPATQITHNESFHEVANEVTSFLTSNQEESTELRGTEI
LNCWNTINTELEETKYNKSHIRDEGVKYKNGLEFLIWOAKMPLNLSHNTYI
ERNKTKNGKRSVDVTFOQAGLSAQVOYADHLRTSMNHLIELTLTKWCREQ
KKNLWKVELSKINPVSMAAIYKPVAVKMGADAPWSECINVDQASVNIHKSMTD
DPKVCYSRPLVTFKPNVSTATFSGOLGTRNEILLTNVTHVETCRPTADHYFFVKWTHY
FKDYFKVMTNNTSIDTLFNLINLIFDINIDFKVELYSETERQASALDLFWER
EYNYTQKGLASREDDNTDLNRDLKOLSEMDADLDIGIKVYNTVTSGLTVFGS
IVGFGVFFNPIGGVTIILLIVVFPVIVSRRTNNWNEAFIKMIYINIDKASQE
NIQPLGEEIKRILLGHQIQQSEHGKSEBEASHKPLGLFOLLGDGLQLLRRTLRP
TFDPSPGNDTSETHQKV"
14184. .17201
/gene="DPOL"
14184. .17201
/gene="DPOL"
/notes="ORF 9"
/codon_start=1
/product="DNA polymerase"

```

```

/protein_id="AAK07928.1"
/db_xref="GI:12802536"
/translation="MEFFNPYISKKTKDKRISCVDPPDKKTNPYCLVYKCFKFTPTGP
GVISITTDPEVLVFOCKEPIITSGHORSYSALKPTPAPIPKDKLFFHYDIIIE
TIYTPORCPDPIQOTDIIPNGTIVLKLKQDQASVCVNVQMVFFVYVLPDGVN
LSVIOQTQNGVNGKCTKESKILKEYPDLSYVYKTIISAPTEINQLVGNL
TSGCVFESNVASRFLIDNKEFGWYCSNPHRISQDRTDLEDFDGLGLQ
FHEBQOQPYITTFIDICIGQGFPCATKDEIYIQLISCIITWCSBPONILLS
LGTCDDPANTVVEFSELMQYFAFFILLRDNIDFVYGNIANFPPYIILDRATOVY
NIPIRFTKVISGTSFVQPMNTGAGFMRSKIKISGLPVIDMTVLDLKLSLSDY
KLANTAUKMGTKQEDVYKBIPLSFSGEAGRARISYCVLSDVLDLKLKFMHIV
EIEISAKMAKIPARRVLTQGOIRVFSCLLEAAKKEHFIPLIPQAKPGGYQAGTVIN
PIGEYNTPLVVDVFLSYPIIOAHNLVSTMIQDNLHLHLKDPDGYTFHLSGP
IHFVKHOKTSLSTLTATLAKRTIRKELANDDDPMKTIIDKQOLAIVKTCNSYV
GFTVAGSLIPCIPAEITVLOQRTMLEKSAVEMITPERLSIVSYPVCPDMSF
RVYIGDTSLEIECRGYPMQSVLNPADKTEITTKALFKDPIKLEAKETQCCLMLTK
KRGIGLSSDKLVKMGVDLIRKTAACSFVQITTSKEILDVLDRDPEVQAAAYLQOAPA
KYVSDGLPVGFVGVLDVNLNSYACATTSRVSSTLSRPFGEYKATITLPHLTV
YKIMSNEELPQIHDRIPYFIKGPQCKSDLAEDPTTVSQNKIPISVEIIFDKUI
HGVANILQCLIFGNSNMTVEILYFNIPYSFT"
17297. .18577
/notes="ORF 10; BORPE1; conserved in other
gamma-herpesviruses"
/codon_start=1
/product="unknown"
/protein_id="AAK07929.1"
/db_xref="GI:12802537"
/translation="MTSTLVEPSIEGQVNICNGNIIILMTNLKDIPIVQGGYGLVKV
FLPLSNQNLNNYISGLVOSIPELEDYAOCTAMLVOCPPGRTPLEIPIHVTEFN
MPFLFVKTQNTIINSLSLAIIFPKPTGCTGICSVISTAPBETIPQSLAHLPS
TDQHLISVANGKLNKHTYMTFRGPPSQPFPMKMTASGRDLODQCVKMGVSD
IKNIMVAVNPGNLVYSVNITPHKPDHIWITLKFTVMAGDLDVLSVLPSTLIS
CGKVNITDSEKILKEGETLNLKTYTRGNESTKAVMFTGTNTNPMVTEFTTW
LPMTPLOVTKNPTNMTIITIKDLAIACVPTYSTLEDROPSPSPSYFNPQDLTITW
EDSMNVATLGENIYSRCHLNKLSNTPSPMDTP"
complement(1848..19345)
/notes="ORF B04; short ORF of immediate early transcript 1"
Query Match 9.3%; Score 847.4; DB 14; Length 108873;
Best Local Similarity 60.2%; Pred. No. 2.6e-212;
Matches 1444; Conservative 0; Mismatches 941; Indels 12; Gaps 2;
Qy 1 ATGGAGATTCGGTCGGATGGACCAACACGTCGAAACCAATCCGGGTTTCATCTTGAC 60
Db 56328 ATGGAAGACGAGTGGGGTGCACTGTATGAGGAGACCAAGTGGGGTTCATTCTAAT 56387
Qy 61 ATGACGTCGATGCCAAAGTCAGGGGTGTCGTGGATCACGTCAGTCGCTGCAATATA 120
Db 56388 ATGACCTCAGATGCCAGGTGAGATCTATCTTGAGAAAGTAAACCTTTATCTGAAA 56447
Qy 121 ACTACACGCCACCGGAAATGGGTGGTACACCTGCCTTCGATCCGGCTGAAGACTCC 180
Db 56448 ACCACAGAGCCCCGGAAATGAATGGGTATGATGCTGAATTTGATCCAAAGTAAATGCC 56507
Qy 181 GGGCCGTTCTTGCCGTTTACCGTTTATCTAAATTACGGGAACTGCTGGTGTGGGAAAAGT 240
Db 56508 GGACCCCTTTTACCTTTTCTGTCTATTGATCACTGSAACAGCTGGTGTGGGAAAAGC 56567
Qy 241 ACCAGCATATCGGCCCTGTACAAAATTTAACTGCCTGATCACGGGCGCCGACCACTA 300
Db 56568 ACAGTATTTCTCCCTCTATCAAAGTGTGAACGTCTTATCACTGGTGCACATACATC 56627
Qy 301 GCGGCACAGAACCTATCGCTCGCTAAGACGTTCTCTCCACGATCTTCAGCGCTTTT 360
Db 56628 CGAGCTCAGAAATTTATCTAAGAGACTGAAACTTACTGCCCCACAAATATTCAATGCC 56687
Qy 361 GCGTTTAAAGACCGCACATCAATATAGCCGTCAGAAAAGCTCATCAGACGGGAGCCGTA 420
Db 56688 GGATTTAAAAGCAGACATATAAATCTTAATCTAGATCACCACCCGGGGACAAATGGACA 56747
Qy 421 TCCATAGAGCAATTCAGCAACAGAGCTATCGAGTAGTATGCGCGGTATATAGTGGACATT 480
Db 56748 AGCATAGATCACTTACAGACATATGAACTCGACAAATATTGGCCCAATCATCTCTGATT 56807

```



QY 481 ATGAAGAGGTTATGCGCAAAAAAACCATGCGATGATCGGACATATATCAACGCGAAT 540  
 Db 56808 TGTAAGGAATTAAGAAAAGAGAAAGTGGCAATATCAAAATATATCTAAATGCTGCT 56867  
 QY 541 TTTGAAACCTCTCGAAGATGACCGGACCGTGTGTTATGACATTCGAATTTATTTGTAATC 600  
 Db 56868 TTTAAATATGTTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 56927  
 QY 601 GACGAGCGCGAACCCTGCTCTTACATATCTACATCTACATCTACATCTACATCTACATCT 660  
 Db 56928 GATGAGCGCGGACATATATCAATATCTATCTATCTATCTATCTATCTATCTATCTATCT 56987  
 QY 661 CTAAACAGCTGGTAAATACCCCTCTTTACCGGAGGCGGCGGTTCCGTTGCAATGATGTC 720  
 Db 56988 TATAATAGCTGGTTAACTCCCTCTATACAGAGCGGAAAGTTCCATGATTTGCTGCT 57047  
 QY 721 GTGCGTTCCCAACGAGCAAAAGCGTTCCAGTCAACTTACAAACGAGGAGCAAAAG 780  
 Db 57048 GTGGGATCCCGACGAGACAGATGCTTCCAGTCCATCTTATATCATATATATATATATAT 57107  
 QY 781 ACGGAGATATGCTGTCGAGAACATCTTAAATCATGATAGGGAAGAGTCTGATCT 840  
 Db 57108 CATCATCTCAGCATGTGCAATGTGTTATCCCTGCTGATCACCATAAGATATGCT 57167  
 QY 841 GAGTACGTACATTTGAGAGAACTGGCGCTGTTTATATAAACAAGAGGCTGACCGAT 900  
 Db 57168 GATTACATAGATCTAAATACAACTGGGCATTTGTTTATCAATACAGAGATGTAATGCT 57227  
 QY 901 CTGCACTTCGACACCTCTCTAAATTTTATAGATATATCTTCCCATCTCTGACGAGTCT 960  
 Db 57228 CCAGAGTTTGGACATATGCTCAAACTTTGGAATATATCTCCCAATACCTGATGATACC 57287  
 QY 961 ATGAGTTAGTACAGATATTTGCTTCTTAAAGTAAGATATGATGATCTTAAATATAC 1020  
 Db 57288 ATGAATTACATAGATATTTATGTTCCCAAGGCAAGATCTCAATCTCTAGATAT 57347  
 QY 1021 ATTTGCTGGACCGGCTCTTTTGTCAATAGCGGTAAAGCGGTATCTTCAACAACTTA 1080  
 Db 57348 ATTTGATGAGCAAGACTGTTTATCCCAATAGGTAAAGCTTACCTGCTCTTTA 57407  
 QY 1081 CACACATGCTAACGTTAGGGGCGATACAGGACACGAAGTCTTTACCTGTCGCGTG 1140  
 Db 57408 CATACTGCTGCTGCAGCTACTAAACAAACCGGGGCAAGCTATTCACATGTCCCAAT 57467  
 QY 1141 GTGTGCGAGTGTGTTGTAAGCGGTTTGAAGGATACAGCGGCGCTCAACCTCACCAAC 1200  
 Db 57468 GTGTGTAATTTTACAGCGGCTTTTCAAAATATTAATCTCACTGTAACCTGATACT 57527  
 QY 1201 CTCACCGTGACCGAGTGGGTGCAAAAAACCTTTTAAAGTTAAGTAATTTTGGAGTTT 1260  
 Db 57528 CTAAACCAACAGATGGTTAATAAGAACTTATTTAGATTAAGTAATTTACTCACAATTT 57587  
 QY 1261 GTGACCGAGCATGTCCATAGTCCGACCGGATCCACTGAAGCTTCGACCCAGGTTACC 1320  
 Db 57588 ATAGACCAAGACATGCAATTTGATCTACAGAAACATCTGAAGATAGTCTTAAGATTACA 57647  
 QY 1321 TTTATCACCAAGTTTGTAAAAACAGCCAGCTGCTTTTAAATGGAAAAACAAAAAATGTC 1380  
 Db 57648 TATGTACAAAATATGTGACGATAGTATGTTTCTCTCAATGTAAGAAAAACAAAAATCT 57707  
 QY 1381 ATATGGGTTTTCAGGGTAGTACTTTGATGTTCAAAAGAAATCTTAGACAGCACTCTTC 1440  
 Db 57708 ACTTGTGGCTTTATGGGCCCATATGAATCTTTTAAAAAGGATATTAAGATGTGAACATTC 57767  
 QY 1441 GTGAGACGATTTTCGCAAGACCGTCCGGAATACGTATACGTTTCTTAAACACGCTACTG 1500  
 Db 57768 ATAGACACCATGACATGATCACTCACTGATTTTGTTCAGTTTATTAATATCTCTCTG 57827  
 QY 1501 TACAACCCCATGACTGCTGTTTCAAGCGGTACGCGGTGACTAGGTCAACAGAGAAATATCTG 1560  
 Db 57828 TATAATGGGATGTATAGTTTCTATATACGGGAGTTAAGGAAAAAACAATAGACTATCTT 57887

QY 1561 CAAGACCTAAATTTGACCCCTCCCG-----GCCGCTCTGCAACCGGCGGTGTA 1611  
 Db 57888 TCAAAACTAAATCAGATACCCATCCCTCAAATATCTCCACCCGGGAACCAAAAAGT 57947  
 QY 1612 GACCTTCAACGCTCGTGAAGAGTTAAACCTGGAAGACACATCTTCTACCACTGTCG 1671  
 Db 57948 GATTACATATCTGATAGTCCGAGAAACAATTGGAAGAGATAGATTTTACTATATGTTG 58007  
 QY 1672 AGTCCCGCCGACCCGCGGTATCACTCCCTCCAGGTTTGTGTCGACGACTACTGCGCC 1731  
 Db 58008 CAGCCCGCCACCAACAGCAATGCTGACCTTTTACCTGTTTGTATGCTGCTATTACAATGCT 58067  
 QY 1732 CTAAGGACGCTGTTCCGCTCCAGAAATAAGGTGCGGTGCGCTGTTTGGCGGAGTTT 1791  
 Db 58068 CTAAGGAATATTTTACTAAGCTTTTCCCTGGCCATAGAAAGCTTTGGACATCTTTT 58127  
 QY 1792 GAGAAGGAACGTTTTCGCGTTTACGGTTAAACATGCTGTAAGGAGCGAGTTGACTTT 1851  
 Db 58128 CTTCAAAAAAATTTGTAACATTTACAATAATATTTGATATCAGAGATGAGTTGAACTC 58187  
 QY 1852 GTCTCCCTTTCAGAACGCTCTCAACGGCTGTTGGCGTTTGCATCGACCGTTGAAATCTAT 1911  
 Db 58188 ACATCTGATGCAACCCACATTTTGGCTCTTGAATATGATCTATCTGTGGAATCATAT 58247  
 QY 1912 AAATTAAGGGGTACACGTTTTCACCGTAGCGTTCCGTCGCTGTCAAGGTTTGCCTACT 1971  
 Db 58248 AAGCTCAAGGATATATCAATCTCCCAAGTTAAATTTTGGAAAGATATGCTGCTGGAGCTA 58307  
 QY 1972 AGTATGACCTCAGGAAGAGATGCTCTCCCTGCTGTCGAGGACTCTAGCGGTTTATC 2031  
 Db 58308 AGTAAGGATCTAAAAACAAATGCCCCAATATATAGTTAGTGTGAAGACTCCCTAGTTTCAIT 58367  
 QY 2032 CGGTGCTTAGAATAACATAAACAAATTCAGCGAAACCATGAGGAGCGGAGCGTTTTC 2091  
 Db 58368 TCTTGCTTAGAATATATGTAATAAATTTACAGAAACCTTAGAGGATGGAGTGTGTTTC 58427  
 QY 2092 CAAGTGTGCTGCGGGGACTATGCGGTGAGTCAAAATTTGCCATGACCATCTGTAAG 2151  
 Db 58428 CATCTTTGATGCGGGTGAATGAGTGTGAGCTTAACTGCGGTATGATTTGTAAG 58487  
 QY 2152 GCACAGGATATGCTGTTGGAGCGAGTAGCCGTAGTATTTGGGTCCCAAGAAACGTCAG 2211  
 Db 58488 GCACAGGATATCTTTGGACAAAGTTGCAATCTCATTTGGAAATCATAAACATCAAA 58547  
 QY 2212 ACAAGCCAGCTGATGATGAGCAATATCAAGGCTGTTTAACTCAAAATTTTGGTCAATGAC 2271  
 Db 58548 CTTAGTCAATGATGATGTTGGGATCTCGAGGCGCCACGAGCCCGAGATGTTTGGTAATAGAT 58607  
 QY 2272 AGCAACCCCTTAAACCCCTCTCAGAGAACCCAGTCGATTAACACCTCCGCAAGCATATA 2331  
 Db 58608 TGCACCCACCTAAACACATATTTCCCTCCACGAGCCAAA---TTCATCTAGTCATATT 58664  
 QY 2332 GTCCGCGCTTCCACACCCCAACACACACCTCATCTACTATAAATAAATAAACAAG 2388  
 Db 58665 GTCAAGGCGCTCCATAACCCCAAAACCTCTGTTTATTATAAGATGAGGAACAG 58721

RESULT 12  
 AF083424  
 LOCUS Ateline herpesvirus 3 complete genome. 108409 bp DNA linear VRL 07-JAN-2000  
 DEFINITION Ateline herpesvirus 3 complete genome.  
 ACCESSION AF083424  
 VERSION AF083424.1 GI:4019231  
 KEYWORDS Ateline herpesvirus 3  
 ORGANISM Ateline herpesvirus 3  
 Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 Gamaherpesvirinae; Rhadinovirus.  
 1 (bases 1 to 108409)  
 Albrecht,J.C.  
 Primary structure of the Herpesvirus ateles genome  
 J. Virol. 74 (2), 1033-1037 (2000)  
 MEDLINE 20091363

```

10623770
REFERENCE
2 (bases 1 to 108409)
AUTHORS
Albrecht, J.-C. and Fleckenstein, B.
TITLE
Direct Submission
JOURNAL
Submitted (11-AUG-1998) Institut fuer klinische und molekulare
Virologie, Universitaet Erlangen-Nuernberg, Schlossgarten 4,
Erlangen, Bayern 91054, Germany
FEATURES
Location/Qualifiers
1..108409
/organism="Ateline herpesvirus 3"
/mol_type="genomic DNA"
/strain="73"
/db_xref="taxon:85618"
complement(join(90..730,1339..1507))
/notes="TCO; similar to Tip and STP-C of herpesvirus
saimiri C-488"
/codon_start=1
/evidence=experimental
/product="tio"
/protein_id="AAC95538.1"
/db_xref="GI:4019240"
/translation="MANEPOHEEGKPPPPPLGDSGEGPPNIPODTPGTPPGPINS
KNEYPPPEPKNKSEPPPGSGNSPPVTMLYKNGDRTKQDVSSEGGNNAPNS
RSPFNKYKRLVDMGENWPPGIPPPQLPRPANLQKQASATSKNGPQIUREAT
EVESQATDQGNHRVKEVKLTVCICLLGILVLLFLMLGFLFLMK"
complement(2144..2264)
/product="HAUR 1"
/product="HAUR 2"
complement(2752..2865)
/product="HAUR 1"
/product="HAUR 2"
3224..6961
/notes="orf03; similar to FGARATs"
/codon_start=1
/protein_id="AAC95537.1"
/db_xref="GI:4019239"
/translation="MLVLFSTVEALLPEELATSAALQPGLOVTHSNPVASGVY
VYKPHNAFSEVLGKFRSPKPTSPAEQQLSLSYSGQLSLOTTSRELVDI
LKLKSLTQGGPKNAIIEKICIRITIIICDINMPTKMQHKLQKLCVYTGILTF
ILQDSIALQAPAAALSTPISDEPLWMCNGLFTLPGYSDKISIFRQGLPKVFE
KNSRLGIVLIDITWTPSSNTIILEGVNVEGVVYFGFTQPTNLMGLVLMALHUR
APLFAACOTGHITSLAQMDMPVQAPQMKMLHNLKICSSMPTIQGFQRPVSDI
ATILQKLYNGSVTALRPTVTKATEASLREKTVLFOLGNYFNRYHLTWDSGK
ETGSLILASFPMDCVKTBGSIIVGMTSTLPIESMKALAAVQCHVSARILLALSALP
QVTSKLAFSNTKTHNKVQLQHPFNLTSLIIATKINLINASEALKTACDMAGCK
FKGILVHRKGVVDDRGDKPYKQDFRFPKPLPKSPINITLQGNINWSELD
LRDTIFKILEHPSGVGKBEIVSHDKLVSGRVARTAIQVWPQLPVSYSIIIVPLOGN
NRDMGNPDYETDIDRFGVFEELAKPGICSSIGESTLIAQDLKIGTMRAITEALLN
LAWNNWNTGSIIOVAITLPTVTAQAATSLQVMDTAKTFCALSVGCTFTASATDGG
ASTVASAIVNTFNAKCMTEPNLKHNDSPILLPTKRYNHFSGVYVQICGKTFIGDLP
QATSPTELKLLQVICTILREDQSLVSGHVDGGLAALAVEMAIISGGGVAVYVPGQE
DTIQUSSETPGIVIVPSKGLYVQVGLNLRNDFQLVKGKICISPLEFAVNSQSDIM
SEPLDSLRQIWRKFSACEADQVHPQVEMHITTPKFCPAGVCFHVSIVYLLPHNS
VPHGLLSIAEAGFQPKLISIDHPYQTSNIYDPTAWGFFIVGASNMQBEENIGIKALK
AQLKSHVALQORDIRIMLAKPEVCAIGAMACELLFYHXTIYGNKPSDPTMYCVQNS
EKPEWSNNTIYIPENTKAFQSLKNSLLPCWTOGTHLRFYHPKPMLEKMAHGMVSS
MFYGNHLASGPAOHPYLTNGSNATAGICADGRHLAHLADPSLCNNICQWPYVPTD
PLKXSPWKTMLDLHLKHWGITVQGSAPPSPRTDDPLRLSLVF"
7470..8552
/notes="orf 04a"
/codon_start=1
/product="complement control protein homolog ccph"
/protein_id="AAC95530.1"
/db_xref="GI:4019232"
/translation="MHTLHFIFILPCVILYFVWSSCKPENEVSLRVYVNTNSGSGY
PNGTTLQVTKRGYIGRQIQITVTCVNGNMTVPNECKRRCSTPADLLNGWYTVGNLY
YGSVITYTCNTGYCLLGSPTSSCLLGDGRVNTWPPPIECCKPPTIANGHTIN
IKSYITVLDATVYSCNDETKLITLTPSSKQCSSETGRWVDEETKCEKFPKQVANG
HVEYKTSNNVOYQYVNIKCDKGRFLQGETTPNCKNGVWFPALPCEKCPAPRGMPH
IDSGEDTSTPSGRNCNQCNTSVSTNIYIITTGHTSHIYFTGKYNKLPQGVLLIIL
TTCGSIITAILITLQVCLHRCRLRMFVP"
9205..12591
/notes="orf 06"
/codon_start=1
/product="major ssDNA binding protein"
/protein_id="AAC95531.1"
/db_xref="GI:4019233"
/translation="MATKATQPIADNLGSRAPVAGCGYIIVYKPEGPPPOEASLLGNR
NVGADAMSLPLLSGLTVEANFENKVAHKIDMTLSVRVSAYHREAIIVFFNNVFTT
FIPGGLDLSCLDARNLFGYTSFTPRHDLKIDIDKLYAFYFTKDSFCMFVSYSVTLH
FKERLYFGLNVLPLISQOQVQINGREAVKIPLYDEDLFSKHEHLPRFYTPSVKSLH
DLSFTSIAQALRIRDESIVRAEKOSIHQDYKAKVANSKEFALQAVKQODASAPW
IDCIAELAISYGLSELEPOEPCALLDVTWPIEDTETEGRIRKAIQDNWAMSVH
VYTHFGSTNSVLYLTINKQTNKSDQNIYNSFYHQHGLSYAAEATQKENGDPAPSG
AVKFGSGSYTLFHLALASSPSPHLLARNCYMQFCQHQKSTNSNTNSVAGYVTVAVS
ELCELQCGKCPAACTIKTFLYRLKDRFPFVLSQRDPYVTVYGVGQNDNDMLGNFAT
FREKEDMQNTQCEKTYWQLIQNVVEKATIGITEGLGSLDITDIOFLKTFREI
DNVVDNEVIFINCLVKNINFRETIKGVHVLHYCNVFWQAPCAFMFLNFYKSVLIA
IIODCLPAMVIEODNPSMMPSEWLKWHYOTIWNFKSCIDIGKVLFGASHKVH
GMDECFINDVSALNGIOLVPKMOVLAKALITVPTIKIKNRIVFSNSMTAIOAG
FIRSSYKSDYIIVTGYMKFLNSHKVIPPDAKISALYLWHTFSKQRIQVLPGISKE
NIELANYIEAGSMHDDMMVLDLIIFITLLIAYAKRLNNTLILATCGTQVYATLOCL
LPTLQVVSATEYFPHVLHQISITSDVDSISRNKQALIVQTLLEKEDATIGKRPYIT
VPLVWNTKTINGNTQIQCGNLGYFMGRGVRNLI.PDSSGFRQNNASLWKRHFIM
PTVANLLKNTLNLLTFEYTRIRKNVQITFDKONLNI.FDNVLIVLVKLGESCEIN
TMDDDLQFYLGDVYIMSDIEWSRFQLLTDGTFWSESVTKILGSRKQEDCNLEFVG
EOLSCAPPOIDBFVQATLSTIAAGKRKIAHILSDIL"
12598..14640
/notes="orf 07"
/codon_start=1
/product="transport protein"
/protein_id="AAC95539.1"
/db_xref="GI:4019241"
/translation="MAQHLAAVSYQIVGYFTLDVSLVTFVDPSPSHINRKFITNNMQKINK
IYLIQIMPELOQNNVESTSLSELOHLNKLMDLIILKLTNYEYFNTYHSLSTSP
CSKHLIVFOFYNCVSVMKCIINDIEIFELKLSVIFYCIRAHDAFYGLNQLNPELV
HLRGISFIPLPDITLSPICLYCLWEHMLPNQGESLOSMTSVNCTVNCQVNLNPELV
QMPENELLQRIIIVDTPCKEPTQCCSINDKIRDETLSLKSHTTIFENTISAPVLES
NLIVWSGTHCTKNTIENTSEIVKLASHETMQNRYKICKSTHFPDKYPSYETI
FCGGIFNSVDETSLKSDCLAFMKRANYQOLIKQNELFVLNLNLTGKEDTSETST
PVPISBKAANVPEOVINDHARKDAYLOKVDKGLSKYACILDTQAVLSNTLSMR
VWCAVDEIVLKLKHLFDRDQISLMDHLEBESVGFNSKYIKLIVYSQKLSSEH
ISSLTQYGLIIGPLSQNSYFPLPNIALHCLDAAAGALPHHKLITIMWIPTIEP
KDMVQYNKPYITISGLNNIOKEVNFTRRELVSILYNALKEKQLQVLSVLAEEK
NCVKLSQPKSGIYLYTEDRSLFIYRNQGVFKDILYLTMLTHHLQKLSVKNHDS"
14627..17050
/notes="orf 08"
/codon_start=1
/product="glycoprotein B"
/protein_id="AAC95532.1"
/db_xref="GI:4019234"
/translation="MTLNRCVLLIVLTFACQTTTPASSDENGKTAIEKEFKYIV
VDLKXPGLTGATITRYVSQPKIPADAGLWGTYKTRTIVNCEIIVEMFARSDAPTYFV
NGSRREGVAILTNKWSLRAVPKYEINLMDKNYQCHNCMQIEVNGLLNSYCDRQGNKT
TALGDTVYSPPCDAENSQPNASDLSQVDENHTVVDYGNRATSQOQGRKIPFAHTLD
YSGVSEAINKTSVCSNVFKQORAIQTHEDSTYHPIANEITAGTSKETLASFS
EYSCLMDSIDITLIDKIGRVNTHVNGTAQYKTEGGMILLVQPLTAILEEAMIEA
TIVTSPLSTAHTSRRTRRRKRDVSAGNSVLLAQIQAIDKLQSGINNVEELI
TWCREQVQRMWYETAKINPTSAIYQKVPKALGDVIVSTECHNDQTSVLIH
KSLKLTNDVQYRPPVTFKFNSSQLFKQLGARNELLSSELVCHQNAEHFFTA
KNETHYKVLHVLTPLTISTLDTFALNLTAFIENIDFKAVELSGYSGERKLANFVD
LETFREYVNYAQISGLRKDFDONSORNRDRIOFSEILADVLGSGKIVNLIASSA
PSLGGVGTGILNFIKPNLGLMPLFLVGAIIIVLIVRTNNMSQAPIRMIYPDIE
KRSSTVPTPEVILKQILLGHNNQOEYKKEZHKASQSPFLKATDAFLRKGAGY
QISTEDK01"
17123..20152
/notes="orf 09"
/codon_start=1
/product="DNA polymerase"
/protein_id="AAC95533.1"
/db_xref="GI:4019235"
/translation="MEFINPLYTKLAGSKMLKSCCTTKQPIICRLVPKCFRNPSPDRGIV
DSCNSFAPPTFYNGSENWYLGNNQSMWHLKRPCKTIEPVEQSITFFHYDIYETIYS
VRVDNVPFKQTDIIPNGTIVKLGRITLDGASVCANFYQKRIYFVKVPEGNTIYLM
KOALNEKIGRCSQITELVKKILSRYPDPEHEVFKVTIISHSYKSYDYLVANGCE
IFETNDAIRFIIIDNNFSTFGWYTKSACFRITNRDSYTDIEPDCYNDLPEHAERT

```

EWPPYNIAMADIEICIGEBGFPFCAKNEGDLIIQISCVFWQVGPDTTNNKMLLSLGTCAA  
VEDVEVFSEADMLHGFTTDFNVETIITGYNISNFDPLVILDRATQIYNIKLSD  
YSRVKSGSI FOVHMKGDMGFWMSYKISIGIIADMVYKCKDKLSUNYKJLDTVA  
DHCTGAKEDSVYKDIPLPWSGPEGRAGIKRGYCVDSVLVAKLLMFKFKTHVEISEIA  
KLAKIPRRVDTCCQIRVFLSCLAAARENYLIPSVNNVNDIGFQATVITPIPGFY  
NNAVLVDPSLASYSTIOAHNLCTYLPHNSLHNPYLKSSDSTFEMLSGSPHFVK  
KHIOISLLKLTIVLGRKRAIKQAQCEQDTQLDKQKALQITCNAYGVFTGV  
ASGLPCISIAETVLTGRRTMLEKSAFTETMPVLQBIVPHKLNHHEHAKFRVIG  
DTSILFVCEGVGVDTVTRGDFLAFETSEKFNAPIKLESEKTCQICLLILAKKRYIG  
ILSNDKLMKGVDLUVTRACKFVQKTSKILNLIKDPVKAAQULSKLLKADYAFRE  
GLDPGLFLKIDILNSENZENTRQSGMDDLTFTESLRPSISYKTEPNPHLTVYKXII  
KRQEBQPIQHDRIYPFVGVKATATCISYMAEDPQATVAYGVNIPITARDLYDFKLTHGVANI  
IOCLFKDSNKTYSVLNVFVSIPELFSYMAEKSDSHSTET"

CDs

THE UNIVERSITY OF CHICAGO

[illegible][illegible]

50

# STIMULI INTENSITY

## STIMULI INTENT

```

/ncore=011.11; similar2 CO kaji 1f22"
/codon_start=1
/protein_id="AAC95535.1"
/db_xref="GI:4019237"
/translation="MAPTPKIRGVWRRGSAKOSKERNVLTWSSVSHSEHFTIITNNO
BIQLPDVILLARCPSEIRVIGQLNFKFSGRYPGPTITSFVYVGHCEVNIKIP
MTI SDCEQLIKFTITFALECI LIPGSKWILPITFLKDLGLYLCLDEPDTSSIMTTS
CMQNGTILSETPPTKLGSEVITLPEPLPMLNAQTKTPNKMKAAVHTVQVECEVN
SIYKGWHLTALQKKEDVNFPEFVILGUTNMKALIAFQNPPIFCPMDWGRQSIIP
IIVYPSIRIPADHAAVVKVNMYSYTIANITCETCFDHFSDCEWKEPSFVY
INVTNINAIITSSGTLKGAEVFLPAPLCKKIIISKHVQTLPSAVTLPGNVTINS
NKLPKLDLITTYK"
complement (23207..24028)
CDS

```

505

complement (23207, .24028)

Query Match 9.0%; Score 812.8; DB 14; Length 108409;  
Best Local Similarity 59.3%; Pred. No. 3.9e-203;  
Matches 1477; Conservative 0; Mismatches 997; Indels 16; Gaps 5;

|    |       |   |       |
|----|-------|---|-------|
| Qy | 1     | ATGGAGAGTTCCGTCCGATGGACCC---AAACAGTCAACCAAAATCCGGGGTTCATCTTGG       | 57    |
| Db | 57799 | ACGGACAGTTTTTTTGGATGAACCGGAATAATGGCAGAACTAAGTCAGAAATTCATACTG        | 57858 |
| Qy | 58    | AACATCAGCTCCGATGCCAAAGTCAGGGTGTGTCGTGATCAGTCACTAGTCCGCTGTCAAAAT     | 117   |
| Db | 57859 | AACATCAGCTCTGATCGAAGAAGTCAGAAATATCGTGGAGACTATAGAAGAAGTGTCAAAAT      | 57918 |
| Qy | 118   | ATAACTACCAGGCCACCGGAANAATGGGTTCGTACGACCTGGCCCTTCGATCCGGCTGAAGAC     | 177   |
| Db | 57919 | ATTACGACTCGCGCTCCAGAAATGAATTTGTATAATAATCAATTTGACCCAGAAAAATGC        | 57978 |
| Qy | 178   | TCCGGGCGGCTTCTGCGGTTTACGGTTTATCTAAATTTACGGGAACCTGCTGGTCTCGGAAA      | 237   |
| Db | 57979 | TCTGGACCTTTGGCTTCATTTACCACTATGTAAATTTACTGGCACTCAGGGGGCTGGAAA        | 58038 |
| Qy | 238   | AGTACCAGCATATCGGCCCTGTACCAAAATTTAAACTGCCTCATCACGGGCGCGACACC         | 297   |
| Db | 58039 | AGTACAAGTATCTCAGCCCTACATCAAAACCTAAACTGCTGATTACTGCGCAACAACT          | 58098 |
| Qy | 298   | ATAGCCGCACAGAAACCTATCGCGTCGCTTAAGACGTTCTGTCCCAAGATCTTCACGCGT        | 357   |
| Db | 58099 | GTTTGCCTTCTCAAAACTTATCCAGGTGCTGTGAAAACCATACTGCGCCCAACAAATTTATAATGCA | 58158 |
| Qy | 358   | TTTGGGCTTTTAAGAGCGCACATCAATATATAGCCGTCAGAAAAAGCTCATCAGACCGGAGCC     | 417   |
| Db | 58159 | TTTGGGTTCAAGAGCAAGCACTAATATATCTACCCAGATCTGTGCTCGCACTACTATG          | 58218 |

Db 59296 ATTTATAATTCCTATATATATTTTCATAATTTATGGCGTTCTTCAAAATAATGAGGCTTAT 59355  
Qy 1558 CTCGACAGACTAAATTTGCAACCCCTCCGGCGGCTCTGGCA-----ACCGGGGTGTA 1611  
Db 59356 TTGCATGATCTAGCAAAATTTAAAGCTACCAAGAAATTTAAACACATCTGTACACAAGCA 59415  
Qy 1612 GACCTTCAAAACGGTTCGTGAAGAGTTAAACCTGGAAGACGACATCTTCTACCACTGTGTC 1671  
Db 59416 GATTTAGTATTTGAAAGGAACACTCATGCTAGAGACGATGTGTTTATCATCTGTT 59475  
Qy 1672 AGTCCCGCGCACCCCGGGGTATCACTCTCTCCAGGTTTGGTTCGACACGTAAGTGGCC 1731  
Db 59476 GTTCTCTCCACCCACAGCTTCATCCGATCTCTTCCATGCTCTAATAAATGTAACACTGCA 59535  
Qy 1732 CTAAGAGAGCTGTTCCTCCAGAAATAAGGTGGCGTGTGCTGGTGTGGCGGGGAGTTT 1791  
Db 59536 CTAAGAGACATATTTATATCTGCTCTAAAGTTGGCAACAGCTGTGTTGTGAGACATTT 59595  
Qy 1792 GAGAAGAAACGTTTTCGGGTTTACGGTTTAAATCATGCTGTAAAGGACGAGTTCACATTT 1851  
Db 59596 TTAGACCATGAGTTTCATCATTTTACAATTAACATGTTAGTAGATTAATAGAAATTT 59655  
Qy 1852 GTCTCCCTTCAGAACGCTTCACGGGCTGTGGCGTTTGGCATCGACCGTTCGAATCGTAT 1911  
Db 59656 ACCTCTACCAATGAGACACTTCATGACTTTTGAATATGCTTCAACTGCTCGAATCTTAC 59715  
Qy 1912 AAAATTAAGGGGTACAGTTTTCACGGTAGCGTTTCGGTGTGCTCAGGGTTTGGCACTC 1971  
Db 59716 AAATTACAGGTTATATCTTTTACCAGTTCATTTTGAAGAGCTCAGACAACGGTTATA 59775  
Qy 1972 AGTGATGACCTCAGAGAGATGCTCCCTCGTGTGCTGACGACCTCAGAGCTTGTATC 2031  
Db 59776 AGCAAGATTTAACAAGATAAAATGCCAAGTATTTATAGTGCAGATTTCTAGCGATTTATA 59835  
Qy 2032 GCCTGCTCAGAGAAATACATAACAAATTCACGAAACCATGAGGACGCGGCTTTTC 2091  
Db 59836 GCATGTTTGAAGAAATGTTAAACAAATGATTGAAATTTAGATGATGAAAGAGCTTT 59895  
Qy 2092 CAAGTGTCTGCGGGGACTATGGGTGAGCTCAAAATTTAGCCATGACCATCGTAAAG 2151  
Db 59896 CACCTCTGTAGCCCGGAGATTATGGCATAAGCTCCAAACTTCCCATGACTATTGTTAAG 59955  
Qy 2152 GCACAGGAATGCTGTGAGCAGTAGCGTGTAGTATTTGGGTGCCACAGAACGTCCTCAG 2211  
Db 59956 GCCCAAGTATATCTCTGTAAAGTTGCTATATGTTTGGAAACCATAAACCTTGAAA 60015  
Qy 2212 ACAAGCACGTTAGTAGCAATATCAAGGGCTGTAACTCAAAATTTATTTGGTCACTGGAC 2271  
Db 60016 ATAGTCACATTTATGTGTGCAATTTCCAGAGCAACCAACCCCAACACATTTGTTAGGAC 60075  
Qy 2272 AGCAACCCCTTAAACCCCTCCTCAGAGAACCTAGTCGATAACACCTCCGCGCAAGCATATA 2331  
Db 60076 CACAAACCTTTAAAA---CTAATGTGTAATCATGCCCATGTGTGTCTTCCAGCACATA 60132  
Qy 2332 GTCCGCGCCCTCCACAAACCAACCAACCTCATCTACTAATAATAATAAAACAAAGCAA 2391  
Db 60133 ATCAAGGCACTTAACAATCCCAACCCCTATTAGTATATTGAATAAATAAATAAAGTT 60192  
Qy 2392 ATGAAACCATTTGTTTATTTTATTCAGTCCCAACCAACCGCCAGGATGTTGTCAATTTCCAC 2451  
Db 60193 AAATTAATTTAAAGTGTAT-TTTAATAAGCACCGGCCAAACATAACAAATTTATGCCGG 60251  
Qy 2452 ACACCGGGGGGTGGCGCCCATAGTTTTCAC 2481  
Db 60252 TAATTGGGGAGTTCCAGTCATTTTGGTC 60281

RESULT 13  
HSGEND 112930 bp DNA linear VRL 25-OCT-2002  
LOCUS Saimirine herpesvirus 2 complete genome.  
DEFINITION X64346  
ACCESSION

VERSION X64346.1 GI:60320  
KEYWORDS alkaline exonuclease; complement control protein homologue; complete genome; cyclin homologue; DHFR gene; dihydrofolate reductase; DNA polymerase; dUTPase; G protein-coupled receptor homologue; glycoprotein; glycoprotein B; glycoprotein H; helicase; HVS-R transactivator; Ig-G gene; integral membrane protein; major capsid protein; major ssDNA-binding protein; mCpH gene; ribonucleotide reductase, large subunit; ribonucleotide reductase, small subunit; saimiri transformation-associated protein; sCpH gene; STP-A11 gene; thymidine kinase; thymidylate synthase; TK gene; transport protein; tyrosine-specific protein kinase; uracil DNA glucosidase.  
SOURCE Saimirine herpesvirus 2  
ORGANISM Saimirine herpesvirus 2  
VIRUSES; dsDNA viruses, no RNA stage; Herpesviridae; Gammaherpesvirinae; Rhadinovirus.  
REFERENCE 1 (bases 1 to 112930)  
AUTHORS Albrecht J.C., Nicholas J., Biller D., Cameron K.R., Biesinger B., Newman C., Wittmann S., Craxton M.A., Coleman H., Fieckenstein B., and Honess R.W.  
TITLE Primary structure of the herpesvirus saimiri genome  
JOURNAL J. Virol. 66 (8), 5047-5058 (1992)  
MEDLINE 92333688  
PUBMED 1321287  
REFERENCE 2 (bases 1 to 112930)  
AUTHORS Albrecht J.  
TITLE Direct Submission  
JOURNAL Submitted (31-JAN-1992) J. Albrecht, Inst. fuer Klinische und Molekulare Virologie, Friedrich-Alexander Uni, Loeschgestrasse 7, D-8520 Erlangen, FRG  
FEATURES  
source 1..112930  
/organism="Saimirine herpesvirus 2"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10381"  
complement (220..714)  
/gene="STP-A11"  
complement (220..714)  
/gene="STP-A11"  
/note="ORF 01"  
/codon\_start=1  
evidence=experimental  
/product="saimiri transformation-associated protein"  
/protein\_id="CAA45623.1"  
/db\_xref="GI:60321"  
/db\_xref="GOA:P18347"  
/db\_xref="SWISS-PROT:P18347"  
translation="MARGLEGDPQENDESNDPPHNTDERSDGDGPTPYLPVTLN  
AGFGPNPFCLLHGVQSGGPRPTALSGAVGLPTSGSRSSSHLSTPVGLSAVY  
SGCGAGSEEHVAVESVLSRSEHQEQDKCTDCSVTILLLLVILLLIILGLMLVIM  
FKXW"  
complement (3969..4532)  
/gene="DHFR"  
complement (3969..4532)  
/gene="DHFR"  
/note="ORF 02"  
/codon\_start=1  
/product="dihydrofolate reductase"  
/protein\_id="CAA45624.1"  
/db\_xref="GI:60322"  
/db\_xref="GOA:P09503"  
/db\_xref="SWISS-PROT:P09503"  
translation="MVOALNCIVAVAGNMIGKQGNLPWRLMDFKHQFQMTTSSV  
PDKNLYMGKTFWFSIEKNRPLKGRINVLSKELKELPHRAHFLAKSLDDALKL  
QPELANKVDVWVWIGSSVYKAMSYPEDLKLFTVRLMQDFECTFFPEFPLEYKLL  
IEYPSVLSNVQEEKSIKTFEYIEKNH"  
complement (6821..10561)  
/note="unnamed protein product; ORF 03  
similarity to ORF 75 and EBV BNRF1"  
/codon\_start=1  
/protein\_id="CAA45625.1"  
/db\_xref="GI:60323"  
/db\_xref="GOA:Q01000"

```
/db_xref="SWISS-PROT:Q01000"
/translation="MLVLFHLYEAQILPHERLATVSFLEHGPLOVTSFNPPTTTTGVV
VYNLSPNAPFLPEVLKLFQSVSPKLPKFKSEQTLSPSYGPHLQSGQTFPSKDLVDV
VYNLSMFTPQGPENAFVERIEYRTHIIICOLDKPKRHRHQLKMLCPYTGLTNF
VYQKRDMDVLDASVASAPLSDPLMTWNGFLNLLIGSYSSQSVFGRGAAPKAFVFE
KNLSRAEIMDYNWPSSTNNLVDGLHAIEKGLIYFGSFTPTSPSTSVFNAAQHWY
VPMFSAQIGHWYQMLPQTLHSLTSLPMEKVLQAHNLKCCSMPIHQFFRLVPPXD
LSHILPQTKIYNGFTSLICLTQINMANEHPLOKTYLYQLGNFTYHKOPLTCQBSG
KEIGSILTAFLDSVKTYPGTIVGMSTSLPVSAMKLAACVQCVCGARLLVSLAP
PQVAKLAPSPSNRNVENKMLKQYFENLLTSDLIILAKNNVIDAQAALTCACVMACG
RFKKLGLLTHLKGTHVNDTVVERPYLILRPRFRPKPPLSSPTISIFLEERINWQDL
NLRDILILKHPSCKEFIVNHTDILISGRVARTSIVGWPQLPVSIDYSLIVPMHFC
NLSVRENPMNIGETIDISHCVFBEAAVAGISCAIGESTILQADLKVGTIRATFEAL
TMSVPPNNINIVTIDITLPTAHVSTYQLVMEATAKTECEALRVSCPTANATEG
GASIVASAIWNTLDVSKITPDLPKFNNSFLLETKEDYSLFGSAQILKEKTYIGBI
PSATSPVLMKMLSLVLTQIKDES VVSGHVDGGLVATVAEMALSGGKGVVYVPHG
EDAIKELCSETPGVVIEVQSGMYVVOQLHSENIPOIIGESSTLSFESTSONLTKL
VHEPLELPSAWRSVSDACERCOHPHRYRMOIATYPKYCPVGRCHFTWIVYLLPN
SVPHGLLALBEAGQPLRLVSLHQPSTKTNVIDEPTVGMFFIVGASNVQDSDVGRAL
TAQKSHVAFORDLETLMAKDPVFSVAGALACELLFYNAIKGYNKPSDPTVCKNS
SKFERSNYSYIPSTKAIAPQSLKNSLIPTWQTGTHLKFYHPKPMLEKMBEGMVS
SNFYGHSLSSGPAQNPYLPNGENAIAGVSGADGRHLALHDPSLCNNLMQWPVPLE
NPPLVSPWKTMFLDLHKWGITVQGASPPSRSDPLRSFV"
join(10912..11773,11774..11966,11967..12013)
/gene="sCCPH"
join(10912..11773,11967..12013)
/gene="sCCPH"
/note="ORF 04b"
/codon_start=1
/evidence=experimental
/product="complement control protein homologue"
/protein_id="CAA45627.1"
/db_xref="GI:60325"
/db_xref="GOA:Q01016"
/db_xref="SWISS-PROT:Q01016"
/translation="MYTLHYICLVSCVIFYFWTLSCPTRNQYVSVKYVNLNTYSGPY
PNGTTLHVTREGYAKRPVQVTVCVNGWNTVPKKCKKSTPDOLLNGRYVTVTGNY
YGSVITYCNSGYSLIGSTTSACLLKRGGRVDWTPRPICDIKKCKPPQIANGHTN
VKDFYILDVTVSCNDETKLTLTGPSKLCSETGVPNGETKCEFIKCLPQVANA
YVEVRKATSMOYLHINVKYKGFMLYGETPNTCNHGVSPAIPECMKISSPKGDMPG
INSNEDNTPSGAECAPGSNYPISS"
10912..11994
/gene="mCCPH"
10912..11994
/gene="mCCPH"
/note="ORF 04a"
/codon_start=1
/evidence=experimental
/product="complement control protein homologue"
/protein_id="CAA45626.1"
/db_xref="GI:60324"
/db_xref="GOA:Q01016"
/db_xref="SWISS-PROT:Q01016"
/translation="MYTLHYICLVSCVIFYFWTLSCPTRNQYVSVKYVNLNTYSGPY
PNGTTLHVTREGYAKRPVQVTVCVNGWNTVPKKCKKSTPDOLLNGRYVTVTGNY
YGSVITYCNSGYSLIGSTTSACLLKRGGRVDWTPRPICDIKKCKPPQIANGHTN
VKDFYILDVTVSCNDETKLTLTGPSKLCSETGVPNGETKCEFIKCLPQVANA
YVEVRKATSMOYLHINVKYKGFMLYGETPNTCNHGVSPAIPECMKISSPKGDMPG
INSNEDNTPSGRIENGCTTSMPQTCTIITARTYTSHYFTGKTYKLPRGLVILL
TTSFIILGILLGVCLHRCVCSGQ"
<10912..11773
/gene="sCCPH"
/number=1
11774..11966
/gene="sCCPH"
/number=1
11967..12013
/gene="sCCPH"
/number=2
complement(11995..12285)
/note="unnamed protein product; ORF 05
KFLF1"
/codon_start=1
```

```
/protein_id="CAA45628.1"
/db_xref="GI:60326"
/db_xref="SWISS-PROT:P24912"
/translation="WTEIIEHLKAKHLQAEENNSNDPVYEEIQLPSEECNRNTT
SLNDSYDVCVPEDMDYDGGAYTRLDASVDYQLTONGDSHYIEIG"
12584..15970
/note="ORF 06
KFRF1"
/codon_start=1
/product="major ssDNA-binding protein"
/protein_id="CAA45629.1"
/db_xref="GI:60327"
/db_xref="GOA:P24910"
/db_xref="SWISS-PROT:P24910"
/translation="VATKTAQPSADNLGSRAPVPCGYIYVTPKEGFPKAEALLGNK
NVGASAMSPLLSGLTVESNFNFNKAHKKIDMTLLVRVSAHREAIYFFNTLDFE
PIFVGLDILGSDARSLLFGYTNFVPRDLRTVDIKDLYAPFSDSCFPMAYVVTGEG
KERLYEGLNVIPIAQLKVOINGRAVKIPLYDDELFSKSHHEHLPFYIPSVSKYLH
DSVFTSIQALIRIDVESVIRASEKSIQDOVKLAKIVNSQFLOSVCODASAFV
IDCIAAEAGLSFLBAPQPCAVLDYTSNPIFETAETEDRIKAIODWAMMSVH
VYTLFSTSVLYLTKITQTSNKSEQVNTYIFMGHGLATYAADATQRENGEPAPSG
APKFSGGTYTLHLALASFSPHLLARNCYQFQCOHKTNTNNTVSVQYVLTAAAS
DLCELCQGTCPASCIHTLFLYRLKDRPPVVLGQRDPYVTVGSGYNDMLGNFAT
FREKDEAVQNAESKRYTYMOLTONVVEKLSVMTGVTGVSELIITDIOSFLKTPRDI
DNVDSSEVKNWNLVKNINPRETIKTTHVHLHYCNVFWQAPCAMFLNLFYKSVLA
IIQDILPANTYEONPESIGMPSWLVKHVQTIWTFKSSCLDRGLVTSRHKIVH
TFMFCDLNDSALSQIVPMQVQVRLAKALLVPTKIKIRIVFNSNSMTIISQSG
FIKSAKDSYIVTGYPMKFLNSLHKVMPFNKISALYILWHTFSKQPLVLPGISRE
NMVELANYETSMKDDHNVLDIIPITLLTYAKVRLNNTILFETCQOTQFYATLOCL
LPTLOTISATYEPHVLIDQSIIVSDHLSIKDKHALTVQTLKEDIAIVGKQRPVIT
TPLVANKYTINGNTQIFCGNLGYPMWGRVDRNLIIPDSTGFRQNNSSYMRHRVFM
TPVAHLKKNLNLNLTREVTIRKNVONIPEDKDNLIIDNVATLELVKGLGDSCE
IITDDLOPLVGEYIINSDIWSRFOIITSGAPWSVENVTKVLGCKNQBECKFEFVGV
BEQUSVPPQIEEFAFOATLILASRKRKIISLSDIDL"
15977..18016
/note="ORF 07
KFRF2"
/codon_start=1
/product="transport protein"
/protein_id="CAA45630.1"
/db_xref="GI:60328"
/db_xref="GOA:P24911"
/db_xref="SWISS-PROT:P24911"
/translation="VAQHLLAAVYQIYGLTLDVSIILTFVDPAPHINMKTVTKTEKIN
IYLIQIMPLHNNQNIESTSLVLOHLHLKIVLETSFSAHLSLDYNNYFENIHS LAP
CSRHKSIVFQYNNCCVVMKCIINDIEIFSKRLSVFYCIRCDALRGINHVIDFLG
HLRGVSPILPDIYLSNIPCIYCLNEHMLPNQGESILPSLMCMVCKVCKOLNPBPI
QGMFENELQORHIIIVESNKQKQONQTCIDNQIRASLSKLNQHTIFENISAPVLES
NLTYSWGAKHKCANVENTSEMAKLLSYEARQNYRKYICKNSLTHFDFKYPPIESI
FCGGIFSSVDVTYKSLKDCSLAFMRANYQOLIKKQNELFVRLNKILOQEDIVSHAA
SAVPLSDKATIVNPDQVLHDAHARKDAYLQKVTDKGLKSLYTCLETQGVLYSLNSMR
```

|                       |       |   |                     |                  |                    |
|-----------------------|-------|---|---------------------|------------------|--------------------|
| Query Match           |       | 8.58;   | Score 770.4;        | DB 14;           | Length 112930;     |
| Best Local Similarity |       | 58.38;  | Pred. No. 6.9e-192; |                  |                    |
| Matches 1429;         |       | Conservative  | 0;                  | Mismatches 1006; | Indels 15; Gaps 4; |
| Qy                    | 34    | GAACCAAATCCGGGGTTTCATCTTGAACATGACGTCGATGCCAAAGTCAGGGGTGTCGTG    | 93                  |                  |                    |
| Db                    | 63007 | GAACTAAGCCCTGAATTCATCTGAAATGATGACTGACATGACATGACATGACATGATGTA    | 63066               |                  |                    |
| Qy                    | 94    | GATCAGCTAGTCGCTGTCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT | 153                 |                  |                    |
| Db                    | 63067 | GAGAAAGATTAGGAAGCTGTCAATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT  | 63126               |                  |                    |
| Qy                    | 154   | CTGGCCCTTCGATCCGGCTGAAGACTCCGGCCGCTTCTTCCGCTTACCGTTTATCTAATT    | 213                 |                  |                    |
| Db                    | 63127 | GACCAATTTGACCCAGAACCAATGCCCAGGAACCTCTGCTTCAATTTACTTGTATGTAATA   | 63186               |                  |                    |
| Qy                    | 214   | ACGGGAATCTGCTGCTGGGAAAGTACCGAGCATATCGGCCCTGTACCAAAATTTAAAC      | 273                 |                  |                    |
| Db                    | 63187 | ACAGGAACTGAGAGCTGGGAAAGTACAGCAATTTTACGCTCTTACCAAAATCTCAAT       | 63246               |                  |                    |
| Qy                    | 274   | TGCCTGATCATCGGGCGGCGACCAACCATAGCGCGACAGAACCTATTCGCTCGCCTAAAGACG | 333                 |                  |                    |

63247 TGCTTATTACTGGTCAACAACAGTAGCATCTCAAAATCTATCAAGATGCTTGAACA 63306  
334 TTCTGTCCACGATCTTCAGCGCTTTTGGCTTTAAGAGCCGACATCAATATAGCCGTC 393  
63307 TACTGCCCTACAATTTTAAATGCTTTGGATTCAGAGTAAGACATTAATATCTCCCA 63366  
394 AGAAAGAGCTCATCAGACCCGAGCGGTATCCATAGACAAATTCAGCAACAGGAGCTATCG 453  
63367 AGATCTGTGCTCGAGACATTAGATCTATTGGAACAAATACAAATTTTGAATATGT 63426  
454 AAGTATTGGCGGTTATAGTGGACATATGAAGAGGTTATGGCGNAACCAATGGC 513  
63427 AAATATTGGGCAATCTGACAAAGCATTTCCAGAAATTTCTAAGAAAGAAATCTAGGC 63486  
514 ATGTAGCGGACTATATCCAAACGGGAATTTGMAACCCCTCTCGAGATGACCGGACCGTGT 573  
63487 CAATACTCAGCATCAGCTTAGCAGCTTTTAAATGCTAGCAAAATGAC---AACTACT 63543  
574 TTATGGATTCCTCAATATATTGTAATCGAGAGCGCGGAAACCCGTCTCTTACATATC 633  
63544 TTGTGGACCAAAATGTAATTTGTTATAGACAAAGCTGGCACCCTTGTCTTCTCACATCTTA 63603  
634 ACCACGCTGCTCTTCTTACTGCTTCTTAACAGCTGGCTAAATACCCCTCTTTACCGC 693  
63604 ACAGAGTTGTGTTTGTGTTATGTTTATPACAGCTGGCTAAACCCCTTTATACAG 63663  
694 CAGGGGCGGTTCCGTGCATAGTATGCTGCTGCTCCCAACGAGCAAAACGCTTCCAG 753  
63664 TCTGGGCGAGTCCATGCAATGCTGTGCTGCTCTCTACTCAAACGATGCAATCAAT 63723  
754 TCAACTTAAACACGCGGAGCAAAAGACGAGATATCTGCTGCGAGACATCTCAACA 813  
63724 TCTACATACAATCACATACAGCAAAAGTAAATATGAGTGTGACACATATTATCA 63783  
814 TTCAATGATGGGAAGAGTCTGATCTGAGTAGTACACTTGGAGAGAACTGGGCGCTG 873  
63784 TTTATTAATGGAAACAAGTTGCTCTGAGTACATCTCCCTCACTAATATTTGGGCTTTA 63843  
874 TTTAATAACAACAGCGCTGCACGATCTCGAGTTCGGACACCTTCTAATAAATTTAGAG 933  
63844 TTCAATAACAATAAAGGTGCTGACCTGAAATTTGGTCACTTACTTAAACACATTTGAA 63903  
934 TATAATCTTCCATTTCTGACGAAGTCATGAGTTACGTAGACAGATTTGTGCTTCTAAA 993  
63904 TATAGCTTTAAATTTCTCTTAAACAATGGAATATATAGATAGATTTGTTGTACAAA 63963  
994 AGTAAGATTATGGAATCTTTAGAAATACATTTGGCTGACCGGACTCTTTTGTACATAGC 1053  
63964 GCTCAGATTTCTTAAACCATTAGAGTTTCTAGGCTGGACAGACTATTTTGTCTCATGT 64023  
1054 GAGTAAAGGGGTATCTAAACAACTACACATGTCTAACGCTAGGGGCGCATACCAGG 1113  
64024 GAAGTTAAGAGTACTTATCTTCACTTCACTAGTACTGATTTACTGMACTAATGTATCA 64083  
1114 GACACGAGCTTTTACCTGCTGCGGTGTGGAGGTGTTGTGAAGCGGTTTGAAGAA 1173  
64084 GGGGCAAGTGTGTTTACTTGGCTTATGTTGTTGTAAGTATTCATTAAGCTTTTAATGAA 64143  
1174 TACAAGCGGCGCTCAACCTCACCACTCACCGTACCGAGTGGGTGACAAAAACCTC 1233  
64144 TATAATCTCATGTATTAATTTACCTCTTTGACAGCTACTGATGTTTATCAAAAAATTTA 64203  
1234 TTTAAGTTAAGTAATTTTCGAGTTTGTGGACGAGACATGTCCATAGTCGCCACGAA 1293  
64204 CACCGTTTAAGCAACTATCCAGTTTATAGACCAAGATATGACTGTCTTATCATCTGAA 64263  
1294 TCCACTGAACTTGCAGCCAGGTACCTTTATACCAAGTTTGTAAAAACGCCAGCTG 1353  
64264 ACTACAGACGCTCAACAAAAAGTAAACATATCTTCAAAAAATGTTAAAAACACATATTT 64323  
1354 TCTTTAATGAAAAACAAAAAATGCATATGTGGGTTTCAGGCTACGTACTTTGAGTTC 1413

64324 TCTCTAACGGGAAAACTAAAAAGTGTGTGTGGTTATGTAGGCATTTATAAAATTTT 64383  
1414 AAAAGAAATCTTAGACACGGAATCTTTCGTGAGACGATTCGCAAGACCGTCCGGAATAC 1473  
64384 AAAAANAATCTAGAAAGTGAATCTTTTATAGATAGTATCTAACGACCAACCTGAATTT 64443  
1474 GTATACGGTTTCTTAAACACGCTACTGTACAAAGCCATGACTCTGTTTCTACGGCTAGCC 1533  
64444 GTTTACAGTTTGTGTGTACAACTATTATAATTTCTTTATATATTTTTCATATATGGC 64503  
1534 GTGACTAGGTACACAGGAGAAATATCTCAAGACCTAAAAATTTGCACCCCTCCCGCGCGCT 1593  
64504 GTGACTAGGAAATAGTCTTAAATAGCTTAGCAAAATTTAAGTTTACCAAGAAAT 64563  
1594 CTGCAACCGGCGCGTGT-----AGACCTTCAACGGTTCGTGAAGAGTTAAACCTCGAA 1647  
64564 CTAACACACTTGTATACACAAACAGACTTAGATATTGAAAGAGAAAGCACTCATGTTAGAA 64623  
1648 GACGACATCTTCTACACGCTGTGAGTCCCGCCGACCGCGGGTATCACCTCCCTCCAG 1707  
64624 GACGATGTTTATCATATGTTTCTCTCCACCACTGCCCTCATCAGCTTCTCTACCT 64683  
1708 GTTTTGTGCGACAGTACTGCGCCCTAAAGAGCGTGTTCGCTCCAGAAATGAAGGTGGCG 1767  
64684 TGCITTAATAGCTGTTATCTGCACTGAAAGACATATTTATATCTCGCTAAAAATTAGCA 64743  
1768 TGTGCTGTTTGGCGGGGAGTTTGAAGAGAACTTTTCCGCGTTTACGGTTAAACATG 1827  
64744 ACAACTTGTGTTTATATAAATTTTATAGTAGAGTTTACATCATCTTACATTTAATATG 64803  
1828 GTCTGTAAAGGACGGAGTTGACTTTTCTCCCTTCAGAAAGTCTCAACGGCTGTGGCG 1887  
64804 CTAGTCAGAGATAATATAGATTTTACATCTACTAATGGAAGACTACATGACTCTTAGAA 64863  
1888 TTGCAATCGACCGTTGAATCGTATAAATTAAGGGGTACAGTTTACGGTAGCGGTC 1947  
64864 TATGCTTCTACAGTTGAATCATCAAAATTAAGGGTATATCTTTTACCGAGTCACTTT 64923  
1948 GGTGCTGTACGGGTTTGGCACTCAGTGATGACCTCAGGAAGAAAGTCCCTCCCTGCTC 2007  
64924 GGAAGATCTCAACTACTGTTTAAAGCAAGACTTCAAGACAAATGCGAAGTATTGTA 64983  
2008 GTGAGAGCTCTAGCGGTTTATCGGTGCTCAGAAATACATAACCAATTAAGCCGAA 2067  
64984 GTGCAAGATTCCAGTGGATTTATAGCTTTTGAAGAAAAATGTCAACAAATGCTAGAA 65043  
2068 ACCATGGAGGACGGAGGCTTTTCCAGTGTGCTGCGGGGACTATGGGGTCAGCTCA 2127  
65044 ACTTTAGATGATGNAAGGCTTTTCACTTTGTAGTGTGAGAGACTATGGAATCAGCTCT 65103  
2128 AATTTAGCCATGACCATCGTAAAGCAAGGATGTGTTGGAGCGGTAGCCGTAGTA 2187  
65104 AAACCTTGCTATGACTATTGTAAGGCGCCCAAGGCACTCTCTTGAATAAGTTGCTATATGT 65163  
2188 TTTGGGTCCTCAAGAACGTCAGACAGCCACGATGTATGTAGCAATATCAAGGCGCTGTT 2247  
65164 TTTATGAACCAATAAANAATTAAGTAGTACATATATGTTGCAATTTCAAGAGCCACA 65223  
2248 AACTCAAAATTTTGGTCAAGCAGCAACCCCTTAAACCCCTCCTCAGAGAACCCAGTC 2307  
65224 AACCTTAACCACTTTGTTATGAGCTGTAATCTTTAAAACTATTAGTCAATGACAC- - - 65280  
2308 GNTAACCTTCGCGCAAGCATATAGTCGCGCCCTCCCAACCCCAACCAACCCCTCATC 2367  
65281 CAGTCTATTTCTCCCAACATATTATTAAAGCTTTAATAACCCCAACACTCTATTAGTC 65340  
2368 TACTAAATAAATAAACAAGCAATGGAACAATTTGTTTATTTAGTCCCAACCCAG 2427  
65341 TATTAAATAAACAATAAATACTTAA- - -AATTAGTTACAAATAATGTTTAAATTAAGCCACG 65397  
2428 GCCAGGATAGTGTCTTTCACACACCGGGGGGTGGCCGCTAGTT 2477  
65398 GCCAAACATAACAGAAATTACCAAGCGGGGAGTCCCAGTCAATTTT 65447







|       |      |             |              |            |               |                                     |       |
|-------|------|-------------|--------------|------------|---------------|-------------------------------------|-------|
| 1294  | TCAC | TGAAGCGTTG  | GACCCAGGTTAC | CTTCTTAT   | CAACAAAGTTTGT | TAAAAACAGCCAGTG                     | 1353  |
| 63973 | ACTA | CAGACACAT   | CAACAAAAGTAA | CATATCT    | TACAAAATATGTT | TAAAAACACATATTT                     | 64032 |
| 1354  | TC   | TTAAATG     | AAAAACAAAAAT | TGCATAT    | GTGGGTTT      | TCAGGTAGCTATTTGAGTTC                | 1413  |
| 64033 | TC   | TC          | CAAGGGAAAC   | TAAAAAGTGT | GTGTGGGTTAT   | GTAGGACATTATAAAAATTTT               | 64092 |
| 1414  | AAAA | GAATCCT     | TAGACAGGAA   | CTTTCGT    | TGGAGACGAT    | TCGCAAGACCGTCCGGATAC                | 1473  |
| 64093 | AAAA | AAATAC      | TGGAAGTGAAT  | CTTTTAT    | TAGATAGTCA    | TGTCAACGACCAACCTGGAATTT             | 64152 |
| 1474  | GTAT | ACGGTTCT    | TAAAAACGCGT  | CTGT       | TACAAACGCGAT  | GTACTCGTTTACGCGTACGGC               | 1533  |
| 64153 | GT   | TTACAGTTTT  | TTTGTGTACAA  | CTTTTAT    | TAAATCTTTTAT  | TATATATTTTTCATATAGGC                | 64212 |
| 1534  | GTG  | ACTAGTGT    | CACAGAGAA    | TATCT      | GCAAGACCT     | TAAAAATTTGCACCCCTCCCGCGCCT          | 1593  |
| 64213 | GT   | ACT         | TGAGAAAAT    | TAGTCTT    | ACTTTAAATGACT | TAGCAAAATTTAAAGTTTACCGAGAA          | 64272 |
| 1594  | CTGG | CAACCGGCGTG | - - - - -    | TAGAC      | CTTCAACCGTTC  | GTGTGAAGAGTTAAACCTCGAA              | 1647  |
| 64273 | CTAA | CCCATTTGT   | TATACAAAC    | AGATTTAG   | TATATTGAAG    | AGAAGCACTCATGTTAGAA                 | 64332 |
| 1648  | GAG  | CACTCTT     | CACAGAGAA    | TATCT      | GCAAGACCT     | TAAAAATTTGCACCTCCCTCCAG             | 1707  |
| 64333 | GAC  | GTGTGTTT    | TATCATATG    | TTCTCT     | CTCCACCACT    | TGCCCTCATGAGCTTCTTGCT               | 64392 |
| 1708  | GT   | TTTGGT      | CGACACG      | TACT       | TGCGCCCT      | TAAAGGACGTGTTCGCTCCAGAAATTAAGGTGCGG | 1767  |
| 64393 | TG   | CTTAAAT     | TAAGCTGT     | TATCT      | GCATGAAG      | ACATATTTATATCTTCGCTTAAATTAGCA       | 64452 |
| 1768  | TG   | TCGCTGTGTT  | TGGCGGAG     | TTTNG      | AGGAAAG       | CGTTTTCGGGTTTACGGTTTAAACATG         | 1827  |
| 64453 | ACA  | ACTTGTGTT   | TAGTAAAT     | TTTTAG     | TATAGAGAG     | TTCCATCATCTTACATTAATATG             | 64512 |
| 1828  | GT   | CGTAAGG     | AGCGAG       | TTTGACT    | TTTGTCTCCCT   | TTCAGAACGTTCTCAACGGGCTGTTCGCG       | 1887  |
| 64513 | CTAG | TCAGAGAT    | AAATATAG     | AAATTTAC   | ATCTACTA      | TCTGGAAGACTACATGACTTTTAGAA          | 64572 |
| 1888  | TTT  | GCATCAG     | CGTTGAA      | TCGTATA    | AAATTAAGGGT   | ACAGTTTTCACGGTAGCGTTC               | 1947  |
| 64573 | TAT  | GTCTT       | CACAGTTGAA   | TCATACAA   | ATTAACGGG     | TTACTTTTTTACCAGTCACTTT              | 64632 |
| 1948  | GGT  | CGCTGT      | CAGGGTTGG    | CACT       | CTCAGT        | GTATGACCTCAGAAAGATGCCCTCCCTGGTC     | 2007  |
| 64633 | GGA  | AGATCT      | CAAACTACT    | GTATT      | TAAGCA        | AAAGACTTACAAGACAAATGCCAAGTATTGTA    | 64692 |
| 2008  | GT   | CAGGACT     | CTTAGCCAT    | TGTTAA     | GGCCTTAGAG    | ATACATAACCAATTCACCGAA               | 2067  |
| 64693 | GT   | CAAGAT      | TCCAGTGG     | ATTTAT     | TAGCATGTT     | TAGAAAAAATGTCAACAATATGCTAGAA        | 64752 |
| 2068  | AC   | CATGGAG     | ACGGGAGCG    | TTTTTCCA   | AGTGTGTCTGT   | CGGGGAGCTATGGGTCAGCTCA              | 2127  |
| 64753 | AA   | TTTAGAT     | TGATG        | AAAAAGC    | TTTCA         | CCCTTGTAGTGTCTGGAGACTATGGCATCAGCTCT | 64812 |
| 2128  | AA   | TTTAG       | CATGACCA     | TGTTAA     | AGGCACAGG     | GAATGTGTGTTGGAGCGAGTAGCCGTAGTA      | 2187  |
| 64813 | AA   | ACTTG       | CCCATGACT    | ATTTG      | TAAAGCCCA     | AGGCACATCTCTTGATAAAGTTGCTATATGT     | 64872 |
| 2188  | TTT  | GGGTCC      | CAACAAG      | ACGTC      | CAGAACCG      | CTGTATGTAGCAATATCAAGGGCTGTT         | 2247  |
| 64873 | TTT  | GGTAA       | CAATAAAAA    | ATTAA      | AGTCA         | CATATATGTTGCAATTTTCAAGAGCCACA       | 64932 |
| 2248  | AA   | CTCAAT      | TATTTTGT     | CTAG       | CAGACG        | CAACCCCTTAAACCTCTCTCAGAGAACCGTCT    | 2307  |
| 64933 | AA   | CCCTAAC     | CACATGTT     | TATG       | GHACTGT       | TAATCTTTTAAAACTATTTAGTCAATGACACA    | 64989 |
| 2308  | GATA | ACACCT      | CCCGCA       | GACATAT    | AGTCCG        | GGCCCTCCACAAACCCAAACCAACCCCTCATC    | 2367  |
| 64990 | CAG  | CTATTTCT    | CTCAAC       | CATATATTA  | AGGCTTTAA     | TAAACCCCAACACTCTATTTAGTC            | 65049 |

QY 6111 CGCGTTTCCATGTCGTTCAAAGGTGAGTACAGATGGATCTCCAGTTAGACTCTGTCTCCCT 6170  
Db 1089 GAAACCGGAGTGTGTACCGAATGCCAATCTGTGCCCTCCAGCTCTCAAAATTTTCATC 1148  
QY 6171 GTTAAACACCTTAACTCGAATGGAACAGCCATACACACCGTGAACCTTCTGTATGTCTCT 6230  
Db 1149 TCCAAATACCGGAATTCGGGAATACACACTCCATGTTTCAGTACATGATGACCTAGGTCTC 1208  
QY 6231 AAAAATAACCCCGAAACAGAGCTAAATACCAATGACTGTCAACCCCTACCCCAAGCACGC 6290  
Db 1209 CCCACCAACCCCATAGGACCCAGCTACAGCTTATCCTCCACTAAATACAGGACAGCTA 1268  
QY 6291 CCCGTTACTATT-----AGACCAAGGTGAGTAAACCAAGCTATCCT 6330  
Db 1269 CCGCGACTCATTAAGCCCGCCAGAAACCAAGTAGTGGGTGGCAATGACAGCTCCCT 1328  
QY 6331 TTTAAAAACCCATACGTCGAGTTTGTAAAGTTAAACAGCT-CGTGTATATTTCAGACGCTG 6389  
Db 1329 TTTAAAAAGTCAACCTTACTCGCAAGGGTAGTCTGTTGTGAGAACTACTGTCCAGGCAGC 1388  
QY 6390 TCATAAATGATACCGACAGACAAATCAGGTAATTAACATTTTTTTTAACTTCAAGATTACGA 6449  
Db 1389 CACAAAAATGGCGCAAGATGACAAAGGTAAAGATCGACC--TTTTTATGTATACTGAAAC 1446  
QY 6450 GGGCGTGTTTACAGTGTGTAGGTGGGAGCATATCTGCCAAATGGGAATGGCTGTGATGAAT 6509  
Db 1447 ATCGGTGTTTACAAATGTTAGTGGGAGCAGAGTTGCCCAAGCTCTACGTCCGAACAGT 1506  
QY 6510 CCACATTAGTCTCGGATCAGGTGCTCTGTGTACTTACTAGTGGCGATGAACCTTAAT 6569  
Db 1507 CCGGTGTACGGGCTCTTATTAAAGTGTTCGTGTACTTGACCAAAGCGCGGAACCTTAGT 1566  
QY 6570 TTGCGGTGTTTAGTCTGTACCAACCAAAAACTGCCTGGCTCGCCTTTTAAAGCTCTA 6629  
Db 1567 TGGGTCTGTACAGTCTGTACCGACCAAAAAAGGATCGGGCGTGTCTTTCAGAGAGTTA 1626  
QY 6630 TGCAGGTTTAAATCGTCTCAATAGCTTGTCTTTAAAAATGCAACTCTGTGATCACTTTA 6689  
Db 1627 GGGACGTGCTGATTTATGTGACAAAGTTTAAACACAGTTAAACTCGGTCTCTGTAGT 6749  
QY 6690 TGACAGTGTCCAAAAAACAAGATTTAAACACAGTTAAACTCGGTCTCTGTAGT 6749  
Db 1687 CGACAGTGTCCAAAAAACAAGTTCAGTGTGACGTTTAAATCTGTATCTCGAAAGT 1746  
QY 6750 CCTCGTATTAACCTCTCAACAGAAAAAATTTTTTAAACCAACTCGCCAGGTACTGAA 6809  
Db 1747 CCTCGTAAATGACAGTTTCTACCAAGAAAAAATTTTTTAAACCGCTGGCCATCCACTGAA 1806  
QY 6810 ACBAGTGTACGGACAGTGTGAAAGTGTCTATCATCTCTTCACTCCCTCTTGAAG 6869  
Db 1807 AGGAGGAGCACAGTCCCGTGTGCGTTGTGTAGGATATCCCTAACTTGGAGCGGAGAC 1866  
QY 6870 TTTTGGTTACACCGACCATGTGAGATAGGCACAGCTCAGGGGGAGGTTCGATCGTGA 6929  
Db 1867 GCGCGGACGCTCCCAAAATGGGAGAGGCACCACTCTGTGCACTCGCGGTCTGGGTT 1926  
QY 6930 GGGGTAACAATTCGTGGCGTGGGGGTACTCCAGTTTTAGGATCTCCGGCAGCACCGGTA 6989  
Db 1927 CTGATTCAGGGGCGCGTGGGGGTATTGGAGAGTCAAAACTCTGGGCAGTCCCTTAA 1986  
QY 6990 CGAGTCTTCTGTAATCTTACGACGCCCGCATCCACCGGGCAGTACTGAGTTAATAA 7049  
Db 1987 TGAGCTCTCTCAAAACCTATGACAGCAGGTCCTAGTGGCAGCATGCCGTTAATA 2046  
QY 7050 GTCCCCGAATTTATCGTTCGTAAATTTTGTGGAGCTCTTTTACAGAGTAAACATGTTC 7109  
Db 2047 CACCCCTTATCTTGTGCTTGCAGTTTGTACAACTGCTCAGGGAATAGCCAAATTCG 2106  
QY 7110 TCCGACATCGGGTACGATGAACCGCTCTCTGTGCTGTGTCATGAGCACCTGGAGTA 7169  
Db 2107 CCCTAGCCGCGGAAACAGGTACGGCTCGCTTTGTGGTGTCTGACCAATATCTGAATGG 2166

QY 7170 TCTTAGACAGGTATAGGGCCCGTTCCACTCTAAGGGTTGCCGTTGGTAGACGGATTGC 7229  
Db 2167 TCTTTGCAAGGTATAGGGTCTTCTCAACGTTTAGAGGGGTACGTGGCAGTCTCGATTGA 2226  
QY 7230 GGGCAATCACCGTCAGTGTGTTCAATTTACCGACTATCATCCGAGTATCATCGGATAGTGA 7289  
Db 2227 GGGTGGCAGCAGAGGTATCTAACTCTCTGAATCTGATCCAGCAGGTAATGAT 2286  
QY 7290 ACTTGTAATATGATTTAAAAAGATGCTTTCCCGTAAATGGCGACTCATGGTCTGTGTTT 7349  
Db 2287 ACCTAAAACAGATGGTTGAACAGGTGATCTTTAAGGGGCTTCTCGATGTCTATTGTAAAA 2346  
QY 7350 A-----ACGTTTCAAGGGTGTAAAAAACTCAGAAAGGTCCTCGCTGGACAACTT 7398  
Db 2347 CTATGACACGCCACTCTCTCTTAGGTTAAGAGCTTCGGCGGTCCTGTGTGAAAGCTT 2406  
QY 7399 TTGCGACTCTCCGAACCGCTCAGGATAAGCTGTATCCACCGGTGCTTCTTTATAACGA 7458  
Db 2407 CGTGGCCTCTCGAGCAACTGAAGGCCAACTCTTACCAAGTGTGTGCTCTCTTTATAAATGA 2466  
QY 7459 AGCATACCTCAAAAAACATGACCCAGCTGACCTAAATAAGCTTCACAGAAAGTCTGTGCG 7518  
Db 2467 CGCATCAAAACATCTACGATCCAGTGCCTTAATAGATGGTGGAAAGATGTGTGAT 2526  
QY 7519 ATGTATTGTAATGGAATGTGCTCCCTGGGTCCCATCTCCGGGTCTATCGGGATCTGAA 7578  
Db 2527 TCGGATTATGAAGAATGTTCCAAAGCTTGGTGGCTATGTGCTGTCTTTACAGACATTAA 2586  
QY 7579 TCTCTTTAACTGCTGCTCTACCGGGTTCGGGGTTAAAAACCGCGCGGCGCCAC 7638  
Db 2587 CATGTTTAACTTTCTGCTTTTTTCTGCTCTCTGAAATGAGGACCAAGAGCGCGCGG 2646  
QY 7639 CTGTAACTCCGTCGCGAGAATCGCCAGGAGATCGTGAAGATTTCTACCGAGAGGCG 7698  
Db 2647 GTACAACTGTCATGTCGAGAGGATCCCAAGGCATTATTCCGATCTCTCAGGAGAGAT 2706  
QY 7699 CTTTGTGTTACCGAAAGATGTTTATAGCTCTGCTGCGAGCGGGTCTTTATACCGCC 7758  
Db 2707 CTTATTTGTCACAGAAAGGCAATTTCTGACAGCGCATGACGCGGGTGAACCTGCTTCC 2766  
QY 7759 CCAGTTGCCAGAGTTCTGACAGAGCTCTACCGGAAATGAAGGCCCAAGTGTTTGGGGCG 7818  
Db 2767 AGCCATATGAAGTACTACAGAAATATACACTGAATGAAGGCCAAATGCTTGGGGCG 2826  
QY 7819 GTGGGTCTCTATATGTTGACGGGCTCTATATGCTATTTGCGCTATTTGCGACTCGTCTCT 7878  
Db 2827 CTGAGGGGACTCTGCTGCAATCGAGGCCCAATATGATATTAACTCTTCCCTACTGAA 2886  
QY 7879 CACTTATATACCTTGGATCGGAGGAAAACTGGAACCTTAGACTCAAGACACTGTGCAA 7938  
Db 2887 GCTCTAACAAGTACGATACCGCGGCTGCTCTCTGAGCAGTCCAGGGCCCTCTGCT 2946  
QY 7939 ACTTGTGTTTCAACCATCTTTTCAAGAAATCTTAGCCCTTATGAGCTACTGCGCAA 7998  
Db 2947 TTTGTTTTTCCAAACCGGTCTACCTTCGAGGATTTATGGCGCGCTGGAGATCATGACCAA 3006  
QY 7999 CGGAGAGATGTTCTCTGACAAATTTTATCCATCACCGGTACGGCCGAGAGAGCGGCC 8058  
Db 3007 GGTTCAGTCTCGCCCTGAAAACTTTTACAGCATCACCGGTTCTGCTGAGAAACCGCGCC 3066  
QY 8059 TGTCTGACTGGAAGTACTAGCGGATGACGTGTCCGGAAAGCAGCTTGTCCCGATTC 8118  
Db 3067 AATTACCACCGCAAGGTCACTGGACTGCTCTATCCAGAAAGCGTCTCTATGCCAGAAATC 3126  
QY 8119 CTTATCTGCGCAGTATCGGAACCGGGTCTTCCCGCACCCCTGGTGTGACCTCAGTAA 8178  
Db 3127 TTTAATTTTGCAATCTTGGAGCCAGGACTGTTGCGGCTTCCATGTTAGACCTCAGCGA 3186  
QY 8179 TGTCTTAGAAATCCAGAAATCATCTCAGCGCCCCACCCCTGAGTCAATTTGTCAAC 8238  
Db 3187 TGTGCTGGCAAAACCCCGCGTTATTCTGAGCGCCCTGCTCCCTGAGCCAGTTTGTATTAG 3246  
QY 8239 AAACAGCACCCCGAGTCTGCTCAGTCACTGAGCATATTATAGCCAAACCCAGGGCGTGT 8298

|    |      |  |         |      |
|----|------|--|---------|------|
| Db | 3247 | CAAAACCCATCCCAACATGCCCGCACACCGTCAGCATCATCCCTTTAACC---      | CATCGGG | 3303 |
| Qy | 8299 | TCCCGGCAATGTTTATGTCACACGTCGGAAGCGGTGCACAGAGCATTCACCAACCGGC | 8358    |      |
| Db | 3304 | TACAGACCGCGCTTTATTAGTACGTGGCAGGCCCGGTACAGATATGGTGTACACAC   | 3363    |      |
| Qy | 8359 | ACAGAC   | 8364    |      |
| Db | 3364 | ATCCAC   | 3369    |      |

Search completed: April 11, 2004, 04:15:49  
 Job time : 22587 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 16:16:38 ; Search time 2057 Seconds  
(without alignments)  
18737.920 Million cell updates/sec

Title: US-09-831-000-1\_COPY\_60966\_70038

Perfect score: 9073

Sequence: 1 atggagagtcctcgatg.....ggggccacgagtcgctga 9073

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score  | Query Match | Length | ID | Description        |
|------------|--------|-------------|--------|----|--------------------|
| 1          | 9073   | 100.0       | 128139 | 6  | AAI64291 RRV genom |
| 2          | 9073   | 100.0       | 133719 | 3  | AAC64754 Macaca mu |
| 3          | 1607.6 | 17.7        | 137507 | 2  | AAV19941 KSHV long |
| 4          | 1002   | 11.0        | 35100  | 2  | AAV73803 KSHV LUR  |
| 5          | 714    | 7.9         | 35100  | 2  | AAV73804 KSHV LUR  |
| 6          | 200.8  | 2.2         | 7361   | 2  | AAV62153 HSV-2 str |
| 7          | 200.8  | 2.2         | 117213 | 2  | AAV62176 HSV-2 str |
| 8          | 200.8  | 2.2         | 154746 | 6  | AAV25519 Human her |
| 9          | 199.2  | 2.2         | 8952   | 2  | AAV62130 HSV-2 str |
| 10         | 154    | 1.7         | 125157 | 5  | Aah74202 Nucleotid |
| 11         | 153.4  | 1.7         | 2871   | 2  | AAT85076 Human cyt |
| 12         | 153.4  | 1.7         | 2871   | 2  | AAT85062 Human cyt |
| 13         | 153.4  | 1.7         | 5290   | 8  | ACC70010 Nucleotid |
| 14         | 153.4  | 1.7         | 229354 | 6  | ABQ74179 Human cyt |
| 15         | 152.4  | 1.7         | 124884 | 5  | Aah74201 Nucleotid |
| 16         | 152.4  | 1.7         | 124884 | 7  | ADAI4878 Human her |
| 17         | 137.6  | 1.5         | 1532   | 4  | AAS63241 Escherich |
| 18         | 129.4  | 1.4         | 681    | 7  | ACF71966 Photorhab |
| 19         | 129.4  | 1.4         | 110000 | 7  | ACF67367 51        |
| 20         | 129.4  | 1.4         | 110000 | 7  | Continuation (52 o |
| 21         | 125.8  | 1.4         | 7316   | 3  | ACF65387_3         |
| 22         | 121    | 1.3         | 657    | 7  | AAC62115 BamHI "G" |
| 23         | 121    | 1.3         | 714    | 7  | ABZ40183 N. gonorr |
|            |        |             |        |    | ABZ40182 N. gonorr |

|      |       |     |        |   |                    |
|------|-------|-----|--------|---|--------------------|
| C 24 | 117.6 | 1.3 | 657    | 5 | AAF94351           |
| C 25 | 117.6 | 1.3 | 657    | 6 | ABK64915           |
| C 26 | 117.6 | 1.3 | 110000 | 2 | AAT42063 Haemophil |
| C 27 | 114.6 | 1.3 | 110000 | 3 | Continuation (12 o |
| C 28 | 114.6 | 1.3 | 349980 | 3 | AAF21609 Neisseria |
| C 29 | 114.6 | 1.3 | 349980 | 3 | AAF21610 Neisseria |
| C 30 | 111   | 1.2 | 4161   | 2 | AAV42547 Kpni-Kpni |
| C 31 | 109.4 | 1.2 | 4161   | 2 | AAV42517 Kpni-Kpni |
| C 32 | 107.4 | 1.2 | 1283   | 4 | AAS09498 Atlantic  |
| C 33 | 107.4 | 1.2 | 1355   | 4 | AAS09499 Atlantic  |
| C 34 | 96    | 1.1 | 4100   | 2 | AAZ32019 Human MET |
| C 35 | 96    | 1.1 | 4100   | 5 | AAC90076 L13855 CD |
| C 36 | 95.4  | 1.1 | 110000 | 6 | ABQ69245_04        |
| C 37 | 95.4  | 1.1 | 110000 | 6 | ABQ67197_03        |
| C 38 | 94    | 1.0 | 696    | 9 | ADD26226 Pseudomon |
| C 39 | 94    | 1.0 | 696    | 9 | ADD26228 Pseudomon |
| C 40 | 91    | 1.0 | 119211 | 4 | AAF28553 Genomic f |
| C 41 | 90.2  | 1.0 | 651    | 4 | Aah52516 S. epider |
| C 42 | 90.2  | 1.0 | 657    | 6 | ABN93020 Staphyloc |
| C 43 | 90.2  | 1.0 | 2975   | 4 | Aah53981 S. epider |
| C 44 | 90.2  | 1.0 | 4114   | 4 | Aah54820 S. epider |
| C 45 | 90    | 1.0 | 1967   | 4 | Aah34344 Human col |

## ALIGNMENTS

RESULT 1  
AAI64291  
ID AAI64291 standard; DNA; 128139 BP.  
XX  
AC AAI64291;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE RRV genome nucleotide sequence.  
XX  
KW RRV, rheus rhadinovirus; Japanese macaque virus; multiple sclerosis;  
KW JMV; cytostatic; antiasthmatic; antiallergic; dermatological; vulnery;  
KW gene therapy; leucopenia; thrombocytopaenia; inflammatory disease;  
KW asthma; allergy; dermatitis; virus; ds.  
XX  
OS Macaca mulatta rhadinovirus 17577.  
XX  
FH Key Location/Qualifiers  
FT CDS 1353..2674  
FT FT /\*tag= a  
FT FT /product= "RRV RI"  
FT FT complement(2692..3258)  
FT FT /\*tag= b  
FT FT /product= "dihydrofolate reductase"  
FT FT /label= RRV ORF2  
FT FT /note= "has\_similarity to Kaposi's sarcoma-associated  
FT FT virus (KSHV) open reading frame (ORF) 2"  
FT FT 3676..5613  
FT FT /\*tag= c  
FT FT /product= "complement binding protein"  
FT FT /label= RRV ORF4  
FT FT /note= "has\_similarity to KSHV ORF4"  
FT FT 5035..9443  
FT FT /\*tag= d  
FT FT /product= "ssDNA binding protein"  
FT FT /label= RRV ORF6  
FT FT /note= "has\_similarity to KSHV ORF6"  
FT FT 9468..11528  
FT FT /\*tag= e  
FT FT /product= "transport protein"  
FT FT /label= RRV ORF7  
FT FT /note= "has\_similarity to KSHV ORF7"  
FT FT 11515..14004  
FT FT /\*tag= f  
FT FT /product= "glycoprotein B"  
FT FT /label= RRV ORF8



| FT | CDS   | FT | CDS                                    |
|----|---|----|--|
| FT | /note= "has similarity to KSHV ORF8"                    | FT | 43491. .44408                          |
| FT | 14122. .17166   | FT | /tag= w                                |
| FT | /tag= g   | FT | /product= "capsid protein"             |
| FT | /product= "DNA polymerase protein"                      | FT | /label= RVV_ORF26                      |
| FT | /label= RVV_ORF9  | FT | /note= "has similarity to KSHV ORF26"  |
| FT | /note= "has similarity to KSHV ORF9"                    | FT | 44433. .45242                          |
| FT | 17261. .18511   | FT | /tag= x                                |
| FT | /tag= h   | FT | /label= RVV_ORF27                      |
| FT | /label= RVV_ORF10                                       | FT | /note= "has similarity to KSHV ORF27"  |
| FT | /note= "has similarity to KSHV ORF10"                   | FT | 45408. .45683                          |
| FT | 18520. .19749   | FT | /tag= y                                |
| FT | /tag= i   | FT | /label= RVV_ORF28                      |
| FT | /label= RVV_ORF11                                       | FT | /note= "has similarity to KSHV ORF28"  |
| FT | /note= "has similarity to KSHV ORF11"                   | FT | complement(45733. .46779)              |
| FT | complement(19921. .20544)                               | FT | /tag= z                                |
| FT | /tag= j   | FT | /label= RVV_ORF29b                     |
| FT | /product= "RVV R2"                                      | FT | /note= "has similarity to KSHV ORF29b" |
| FT | /label= RVV_ORF   | FT | 46905. .47135                          |
| FT | /note= "has similarity to KSHV interleukin (IL)-6 gene" | FT | /tag= aa                               |
| FT | complement(20777. .21778)                               | FT | /label= RVV_ORF30                      |
| FT | /tag= k   | FT | /note= "has similarity to KSHV ORF30"  |
| FT | /product= "thymidylate synthase"                        | FT | 47093. .47746                          |
| FT | /label= RVV_ORF70                                       | FT | /tag= ab                               |
| FT | /note= "has similarity to KSHV ORF70"                   | FT | /label= RVV_ORF31                      |
| FT | complement(22245. .22592)                               | FT | /note= "has similarity to KSHV ORF31"  |
| FT | /tag= l   | FT | 47683. .49077                          |
| FT | /product= "RVV R3"                                      | FT | /tag= ac                               |
| FT | /note= "has similarity to KSHV K4 viral MIP gene"       | FT | /label= RVV_ORF32                      |
| FT | 26846. .27409   | FT | /note= "has similarity to KSHV ORF32"  |
| FT | /tag= m   | FT | 49049. .50059                          |
| FT | /product= "Bcl2-homologue"                              | FT | /tag= ad                               |
| FT | /label= RVV_ORF16                                       | FT | /label= RVV_ORF33                      |
| FT | /note= "has similarity to KSHV ORF16"                   | FT | /note= "has similarity to KSHV ORF33"  |
| FT | complement(27515. .29125)                               | FT | complement(49977. .50960)              |
| FT | /tag= n   | FT | /tag= ae                               |
| FT | /label= RVV_ORF17                                       | FT | /label= RVV_ORF29a                     |
| FT | /note= "has similarity to KSHV ORF17"                   | FT | /note= "has similarity to KSHV ORF29a" |
| FT | 28998. .29897   | FT | 50959. .51942                          |
| FT | /tag= o   | FT | /tag= af                               |
| FT | /label= RVV_ORF18                                       | FT | /label= RVV_ORF34                      |
| FT | /note= "has similarity to KSHV ORF18"                   | FT | /note= "has similarity to KSHV ORF34"  |
| FT | complement(29905. .31548)                               | FT | 51923. .52372                          |
| FT | /tag= p   | FT | /tag= ag                               |
| FT | /product= "tegument protein"                            | FT | /label= RVV_ORF35                      |
| FT | /label= RVV_ORF19                                       | FT | /note= "has similarity to KSHV ORF35"  |
| FT | /note= "has similarity to KSHV ORF19"                   | FT | 52278. .53585                          |
| FT | complement(31043. .32095)                               | FT | /tag= ah                               |
| FT | /tag= q   | FT | /product= "kinase"                     |
| FT | /label= RVV_ORF20                                       | FT | /label= RVV_ORF36                      |
| FT | /note= "has similarity to KSHV ORF20"                   | FT | /note= "has similarity to KSHV ORF36"  |
| FT | 32034. .33767   | FT | 53566. .55008                          |
| FT | /tag= r   | FT | /tag= ai                               |
| FT | /product= "thymidine kinase"                            | FT | /product= "alkaline exonuclease"       |
| FT | /label= RVV_ORF21                                       | FT | /label= RVV_ORF37                      |
| FT | /note= "has similarity to KSHV ORF21"                   | FT | /note= "has similarity to KSHV ORF37"  |
| FT | 33754. .35868   | FT | 54963. .55172                          |
| FT | /tag= s   | FT | /tag= aj                               |
| FT | /product= "glycoprotein H"                              | FT | /label= RVV_ORF38                      |
| FT | /label= RVV_ORF22                                       | FT | /note= "has similarity to KSHV ORF38"  |
| FT | /note= "has similarity to KSHV ORF22"                   | FT | complement(55255. .56391)              |
| FT | complement(35865. .37073)                               | FT | /tag= ak                               |
| FT | /tag= t   | FT | /product= "glycoprotein M"             |
| FT | /label= RVV_ORF23                                       | FT | /label= RVV_ORF39                      |
| FT | /note= "has similarity to KSHV ORF23"                   | FT | /note= "has similarity to KSHV ORF39"  |
| FT | complement(37123. .39321)                               | FT | 56526. .57932                          |
| FT | /tag= u   | FT | /tag= al                               |
| FT | /label= RVV_ORF24                                       | FT | /product= "helicase/primase"           |
| FT | /note= "has similarity to KSHV ORF24"                   | FT | /label= RVV_ORF40                      |
| FT | 39323. .43459   | FT | /note= "has similarity to KSHV ORF40"  |
| FT | /tag= v   | FT | 57917. .58528                          |
| FT | /product= "major capsid protein"                        | FT | /tag= am                               |
| FT | /label= RVV_ORF25                                       | FT | /product= "helicase/primase"           |
| FT | /note= "has similarity to KSHV ORF25"                   | FT | /label= RVV_ORF41                      |

|       |   |       |
|-------|---|-------|
| 56226 | GAGTACGTACACTTTGGAGAGAACTGGCGCGCTGTTTATAAACAACAAGCGCTGCACCGAT     | 56288 |
| 901   | CTGCAGTTCCGACACCTTTCTAAAAATTTTATAGATATAAATCTTCCCATTTCTCTGACGAAGTC | 960   |
| 56286 | CTGCAGTTCCGACACCTTTCTAAAAATTTTATAGATATAAATCTTCCCATTTCTCTGACGAAGTC | 56345 |
| 961   | ATGAGTTACGTAGACAGATTGTGCTTAAAGTAAGATTATGATCTCTTTAGATAATC          | 1020  |
| 56346 | ATGAGTTACGTAGACAGATTGTGCTTAAAGTAAGATTATGATCTCTTTAGATAATC          | 56405 |
| 1021  | ATTGGCTGGACCCGACTCTTTTGTGCATAGCGAGGTAAGCGGTATCTTAAACAAACCTTA      | 1080  |
| 56406 | ATTGGCTGGACCCGACTCTTTTGTGCATAGCGAGGTAAGCGGTATCTTAAACAAACCTTA      | 56465 |
| 1081  | CACATGTCATAAGCTTAGGGGGCGATACCAAGGACACAGAGCTCTTTTACCTGTCCTGGTG     | 1140  |
| 56466 | CACATGTCATAAGCTTAGGGGGCGATACCAAGGACACAGAGCTCTTTTACCTGTCCTGGTG     | 56525 |
| 1141  | GTGTCGAGGTGTTGTGTAAGCCGTTTGTGAGGATACAAAGCGGCGCTCAACCTCACCAAC      | 1200  |
| 56526 | GTGTCGAGGTGTTGTGTAAGCCGTTTGTGAGGATACAAAGCGGCGCTCAACCTCACCAAC      | 56585 |
| 1201  | CTCACGCTGACCGAGTGGGTGACAAAAAATCTCTTTAAGTTAAGTAATATTTCGACGTTT      | 1260  |
| 56586 | CTCACGCTGACCGAGTGGGTGACAAAAAATCTCTTTAAGTTAAGTAATATTTCGACGTTT      | 56645 |
| 1261  | GTGGACACGAGCATGCTCCATAGTCGCCACGGAATCCACTGAAACGTTCCGACCGAGTTACC    | 1320  |
| 56646 | GTGGACACGAGCATGCTCCATAGTCGCCACGGAATCCACTGAAACGTTCCGACCGAGTTACC    | 56705 |
| 1321  | TTTATCACCAAGTTTGTAAAAACAGCCACGCTGCTTTAAATGGAATAACAAAAATGC         | 1380  |
| 56706 | TTTATCACCAAGTTTGTAAAAACAGCCACGCTGCTTTAAATGGAATAACAAAAATGC         | 56765 |
| 1381  | ATATGTGGGTTTCAGGGTACGTACTTTGAGTTCAAAAGAAATCCCTAGACAGCGAACTCTTC    | 1440  |
| 56766 | ATATGTGGGTTTCAGGGTACGTACTTTGAGTTCAAAAGAAATCCCTAGACAGCGAACTCTTC    | 56825 |
| 1441  | GTGGAGACGCAATTCGCAAGACCGTCCGGAATACGTTATACGTTTCTTAAACACGCTACTG     | 1500  |
| 56826 | GTGGAGACGCAATTCGCAAGACCGTCCGGAATACGTTATACGTTTCTTAAACACGCTACTG     | 56885 |
| 1501  | TACAACGCACTGCTACTGCTTTTACGCGTACCGGCGTACTAGGTACACAGAGAAATATCTG     | 1560  |
| 56886 | TACAACGCACTGCTACTGCTTTTACGCGTACCGGCGTACTAGGTACACAGAGAAATATCTG     | 56945 |
| 1561  | CAAGACCTAAAAATTTGCAACCCCTCCCGCCGCTCTGGCAACCGGCGGTGTAGACCTTCAA     | 1620  |
| 56946 | CAAGACCTAAAAATTTGCAACCCCTCCCGCCGCTCTGGCAACCGGCGGTGTAGACCTTCAA     | 57005 |
| 1621  | ACGGTTCGTGAAGATTAAACCTGGAAGACACATCTTCTACACGCTGTGACGCTCCCGCG       | 1680  |
| 57006 | ACGGTTCGTGAAGATTAAACCTGGAAGACACATCTTCTACACGCTGTGACGCTCCCGCG       | 57065 |
| 1681  | CCACCCGCGGGTATCACTCCCTCCAGGTTTGGTTCGACAGTACTGCGCCCTTAAGGAC        | 1740  |
| 57066 | CCACCCGCGGGTATCACTCCCTCCAGGTTTGGTTCGACAGTACTGCGCCCTTAAGGAC        | 57125 |
| 1741  | GTGTTCCGCTCCAGATAAAGGTGGCGTGTGCTGTTTGGCGGGAGTTTGAAGAAGAA          | 1800  |
| 57126 | GTGTTCCGCTCCAGATAAAGGTGGCGTGTGCTGTTTGGCGGGAGTTTGAAGAAGAA          | 57185 |
| 1801  | ACGTTTTCCCGGTTACCGTTTAAATGTCGTAAGGACGGAAGTTGACTTTGTCTCCCT         | 1860  |
| 57186 | ACGTTTTCCCGGTTACCGTTTAAATGTCGTAAGGACGGAAGTTGACTTTGTCTCCCT         | 57245 |
| 1861  | TCGAACGCTCAACGGGCTGTGGCTTTCATCGACCGTTGAAATCTGATAAAATTAAG          | 1920  |
| 57246 | TCGAACGCTCAACGGGCTGTGGCTTTCATCGACCGTTGAAATCTGATAAAATTAAG          | 57305 |
| 1921  | GGGTACACGTTTTTACCGGTAGCGTTTCGGTCTGCTGAGGGTTTGCCACTCAGTGTATGAC     | 1980  |
| 57306 | GGGTACACGTTTTTACCGGTAGCGTTTCGGTCTGCTGAGGGTTTGCCACTCAGTGTATGAC     | 57365 |

1981 CTTCAGGAAGAAGATGCCCTCCCTGGTGGTGGAGACTCTAGCGGTTTTATCGGTGCTTA 2040  
57366 CTTCAGGAAGAAGATGCCCTCCCTGGTGGTGGAGACTCTAGCGGTTTTATCGGTGCTTA 57425  
2041 GAGAAATAACATAACCAAAATGACCGAAACCAATGAGGACGGAGCGTTTCCAAAGTGTGC 2100  
57426 GAGAAATAACATAACCAAAATGACCGAAACCAATGAGGACGGAGCGTTTCCAAAGTGTGC 57485  
2101 TGTGGGGGGGACTATGGGTGAGTCACTCAATTTAGCCATGACCATCGTAAAGGACACAGGA 2160  
57486 TGTGGGGGGGACTATGGGTGAGTCACTCAATTTAGCCATGACCATCGTAAAGGACACAGGA 57545  
2161 ATGTGCTGGAGCGGAGTACCGGTAGTATTTGGGTCCCAAGAAACGTCACAGCAAGCCAC 2220  
57546 ATGTGCTGGAGCGGAGTACCGGTAGTATTTGGGTCCCAAGAAACGTCACAGCAAGCCAC 57605  
2221 GTGTATGTAGCAATATCAAGGGCTGTAACTCAATTTATTTGGTCATGAGCAGCAACCC 2280  
57606 GTGTATGTAGCAATATCAAGGGCTGTAACTCAATTTATTTGGTCATGAGCAGCAACCC 57665  
2281 CTTAAACCCCTCTCAGAGAACCGTGTGATTAACACCTCCGCAAGCATATAGTCCGCGCC 2340  
57666 CTTAAACCCCTCTCAGAGAACCGTGTGATTAACACCTCCGCAAGCATATAGTCCGCGCC 57725  
2341 CTCACCAACCCCAACCAACCTCTACTTAATAATAATAAACAAGCAAAATGGAAC 2400  
57726 CTCACCAACCCCAACCAACCTCTACTTAATAATAATAAACAAGCAAAATGGAAC 57785  
2401 ATGTGTTTTTATTCAGTCCCAACCAACCGCCACCGATAGTGTCAATTTCCACACACCGGG 2460  
57786 ATGTGTTTTTATTCAGTCCCAACCAACCGCCACCGATAGTGTCAATTTCCACACACCGGG 57845  
2461 GGGTGGCGCCATAGTTTGAAGACCAAGAACCGCGGCCCTCTGGGTTAGAAAGTAGCC 2520  
57846 GGGTGGCGCCATAGTTTGAAGACCAAGAACCGCGGCCCTCTGGGTTAGAAAGTAGCC 57905  
2521 CGCCCCCGGTGTGAGTGCATACGTCTTTAGGGCCCTGGTGTATACCAACATAATCA 2580  
57906 CGCCCCCGGTGTGAGTGCATACGTCTTTAGGGCCCTGGTGTATACCAACATAATCA 57965  
2581 CGTGTGTATCAAGTTCAGTCCCTCGGTGAGCGCGCCCTCTGGAGCGGCGGACGAC 2640  
57966 CGTGTGTATCAAGTTCAGTCCCTCGGTGAGCGCGCCCTCTGGAGCGGCGGACGAC 58025  
2641 CGCGCTGTGTAACGTGTATCTGCATGTCGTGCTGTCTGAAGATGAACACAGTCAATCA 2700  
58026 CGCGCTGTGTAACGTGTATCTGCATGTCGTGCTGTCTGAAGATGAACACAGTCAATCA 58085  
2701 CATCAGTGGCGCGGCGCCCTCATCGGCTGTGATTCGTCCTCATCAACAGCGCTC 2760  
58086 CATCAGTGGCGCGGCGCCCTCATCGGCTGTGATTCGTCCTCATCAACAGCGCTC 58145  
2761 GGTGTCCACTTCGTGTGTCACTGAAATAATATGTTCTGTCTGTGCAAGTTGATCCAG 2820  
58146 GGTGTCCACTTCGTGTGTCACTGAAATAATATGTTCTGTCTGTGCAAGTTGATCCAG 58205  
2821 CAACCAACAGGTTCTTCTCATCGGAAGCCATCCGCGCTCTGCGGTGATCTAGATATAT 2880  
58206 CAACCAACAGGTTCTTCTCATCGGAAGCCATCCGCGCTCTGCGGTGATCTAGATATAT 58265  
2881 CACTGTACAGCGCGGCGGAGCCCGGTGGCGAAAGCCGTGAGTCTCTCGCCCACTTTCA 2940  
58266 CACTGTACAGCGCGGCGGAGCCCGGTGGCGAAAGCCGTGAGTCTCTCGCCCACTTTCA 58325  
2941 CAGTGTCTGTGTGTATCAATAAATAATGCTCTGCTGTCTGTATTCAGATGACTGG 3000  
58326 CAGTGTCTGTGTGTATCAATAAATAATGCTCTGCTGTCTGTATTCAGATGACTGG 58385  
3001 GTGCGCGGTGTGGCCAGTAAACCAACCGGCGCGGACCAACCAACCGCGGTGAGTG 3060  
58386 GTGCGCGGTGTGGCCAGTAAACCAACCGGCGCGGACCAACCAACCGCGGTGAGTG 58445

3061 GCGCGGGCGCCCTCGGTGGTGGTGGCACTGTGCGAAGTAATGACCGGTAGGTAGCGGC 3120  
58446 GCGCGGGCGCCCTCGGTGGTGGTGGCACTGTGCGAAGTAATGACCGGTAGGTAGCGGC 58505  
3121 CGCTCTGTTTAAAGGACTCACCGTCCCGTGGTGGGATCTATTTCTAGGTGGGAGGAGA 3180  
58506 CGCTCTGTTTAAAGGACTCACCGTCCCGTGGTGGGATCTATTTCTAGGTGGGAGGAGA 58565  
3181 CGCGCGGTTCCTCTCTTCATCCACATCACTGTAGTGGTGGTCTCCGAATGAGTCCG 3240  
58566 CGCGCGGTTCCTCTCTTCATCCACATCACTGTAGTGGTGGTCTCCGAATGAGTCCG 58625  
3241 TTTGCTCTGGCGTTTTTTGGAGTAGTGGGAGGGAATCTTGGTCAAAAAACATCTCTCAGGTG 3300  
58626 TTTGCTCTGGCGTTTTTTGGAGTAGTGGGAGGGAATCTTGGTCAAAAAACATCTCTCAGGTG 58685  
3301 CAGTCATGTCAATCACTGGGGGGCTCCGTACCGGGCGGTACGGCAACCGCGTACGG 3360  
58686 CAGTCATGTCAATCACTGGGGGGCTCCGTACCGGGCGGTACGGCAACCGCGTACGG 58745  
3361 ACCCGGGGAACAAAAAACGTCCTTCTCTTTCTCTAGTGGTGGGGAATCGGACGCA 3420  
58746 ACCCGGGGAACAAAAAACGTCCTTCTCTTTCTCTAGTGGTGGGGAATCGGACGCA 58805  
3421 TCCTGGAGTGGCGGGGGCTCTTCACGGTGGATAGACATAGCATGGCGGAA 3480  
58806 TCCTGGAGTGGCGGGGGCTCTTCACGGTGGATAGACATAGCATGGCGGAA 58865  
3481 CGTTCACCTGTAAACCGACTGCTATCCCGCATCACTTAATATCCAGTCAACGCGACC 3540  
58866 CGTTCACCTGTAAACCGACTGCTATCCCGCATCACTTAATATCCAGTCAACGCGACC 58925  
3541 CCTCCGATTCGAACTAGATAGTCACTTCGCAAGTTTAAATAGTTCGAGCTAGAACTG 3600  
58926 CCTCCGATTCGAACTAGATAGTCACTTCGCAAGTTTAAATAGTTCGAGCTAGAACTG 58985  
3601 CGGCAGGAGAACCGGTTCGCGTGGCGGTAGCGAGGAGACGGATGGCGGCTT 3660  
58986 CGGCAGGAGAACCGGTTCGCGTGGCGGTAGCGAGGAGACGGATGGCGGCTT 59045  
3661 GAGGACGAGATGCTCTGCTCCGTTTATCAGCGTGGCTGTCTATAGCTTTCTTCCCA 3720  
59046 GAGGACGAGATGCTCTGCTCCGTTTATCAGCGTGGCTGTCTATAGCTTTCTTCCCA 59105  
3721 CAGCATAAAAACGCACTCTGAAGCTTGGCAGAAAGCGAGCTTATTTATGTAGTGTGAA 3780  
59106 CAGCATAAAAACGCACTCTGAAGCTTGGCAGAAAGCGAGCTTATTTATGTAGTGTGAA 59165  
3781 CCAATCCAGCAAGTGGCGTGTGACCCCGCTTCCCTCTTCCACCGTCAAGATGCT 3840  
59166 CCAATCCAGCAAGTGGCGTGTGACCCCGCTTCCCTCTTCCACCGTCAAGATGCT 59225  
3841 GTTTAAAGCAGAACCTCCCGTTTTGCCAAACAGTCTAAGCACCCGTGAGAGGAGCGT 3900  
59226 GTTTAAAGCAGAACCTCCCGTTTTGCCAAACAGTCTAAGCACCCGTGAGAGGAGCGT 59285  
3901 GAACCCAGGTACAGTATGGCTATCTTTTAAATAATTTTGGAGCTTGGAGGTATCT 3960  
59286 GAACCCAGGTACAGTATGGCTATCTTTTAAATAATTTTGGAGCTTGGAGGTATCT 59345  
3961 GTAGTCCGGAGCAACGCTAAGGCTAGTCCGTTGGCTTGAACCGCATGTAGGATCTCTG 4020  
59346 GTAGTCCGGAGCAACGCTAAGGCTAGTCCGTTGGCTTGAACCGCATGTAGGATCTCTG 59405  
4021 GCCAAGGATCAGCACTTTAATATCTCCGGTTGCGCAACAAATAGCACCAACATATCT 4080  
59406 GCCAAGGATCAGCACTTTAATATCTCCGGTTGCGCAACAAATAGCACCAACATATCT 59465  
4081 ATCTATCGGTGGATAAATCAACGTTTACGTTGTCTATATCAATAACCGCTTTCAGCAGAC 4140  
59466 ATCTATCGGTGGATAAATCAACGTTTACGTTGTCTATATCAATAACCGCTTTCAGCAGAC 59525  
4141 CGCGAGCTTCTGTCTTTTAAATAATGAGAAAGATTTTAAATAATTCAGGCTAGTCCGTAAG 4200



|    |       |                         |  |       |    |       |                           |   |       |
|----|-------|-------------------------|--|-------|----|-------|---------------------------|---|-------|
| Qy | 6361  | AAACAGCTCGTGTATATTTT    | CAGACGCTGTGCATATAAATGATACCGACGCAATCAGGTA | 6420  | Qy | 7441  | GGTGCTCTTATTAACGAGCATACGT | CAAAAAACATGACCCAGTGCATTAATAAGCT         | 7500  |
| Db | 61746 | AAACAGCTCGTGTATATTTT    | CAGACGCTGTGCATATAAATGATACCGACGCAATCAGGTA | 61805 | Db | 62826 | GGTGCTCTTATTAACGAGCATACGT | CAAAAAACATGACCCAGTGCATTAATAAGCT         | 62885 |
| Qy | 6421  | ATTAACACTTTTATTAACAGATT | CAGAGTTCAGGAGCGGTGTACAGTGGTGTAGTGGAGCA   | 6480  | Qy | 7501  | CACAGAGAAAGTCTGTGATGAT    | TATTTAAATGGAATGTCCTCCCTGGGTCCCATCTCCGG  | 7560  |
| Db | 61806 | ATTAACACTTTTATTAACAGATT | TCAGGAGGGGTGTACAGTGGTGTAGTGGAGCA         | 61865 | Db | 62886 | CACAGAGAAAGTCTGTGATGAT    | TATTTAAATGGAATGTCCTCCCTGGGTCCCATCTCCGG  | 62945 |
| Qy | 6481  | TATCTGCCAATGGGAATGGCT   | GATGAATCCACACTTAGTCTCGGATCAGTGTCTGT      | 6540  | Qy | 7561  | GCTCATCGCGATCTGAATCTCT    | TAACTTTAACTTTGTCCTCTACCGGGGTTCCCGGGTTAA | 7620  |
| Db | 61866 | TATCTGCCAATGGGAATGGCT   | GATGAATCCACACTTAGTCTCGGATCAGTGTCTGT      | 61925 | Db | 62946 | GCTCATCGCGATCTGAATCTCT    | TAACTTTAACTTTGTCCTCTACCGGGGTTCCCGGGTTAA | 63005 |
| Qy | 6541  | GTACTTTACTAGTCCGATGAAC  | CTAAATTTGGCGTGTAGTCTGACCAACCAAAAA        | 6600  | Qy | 7621  | AACCCGGGGGCCCCACCTGTAA    | CGTCCGTCGCGAGAATGCGCCAGGGCATCTGTGAG     | 7680  |
| Db | 61926 | GTACTTTACTAGTCCGATGAAC  | CTAAATTTGGCGTGTAGTCTGACCAACCAAAAA        | 61985 | Db | 63006 | AACCCGGGGGCCCCACCTGTAA    | CGTCCGTCGCGAGAATGCGCCAGGGCATCTGTGAG     | 63065 |
| Qy | 6601  | CTGCTGGCCTCGCTTTTAAAG   | CTCTATGCAAGCTTTAATCGTCTCCAATAGCTTGT      | 6660  | Qy | 7681  | AATTTCTCACGAGAGGCGCTTT    | GTGTACCGAAGAATGTTTATAGCTGTGCTGCAG       | 7740  |
| Db | 61986 | CTGCTGGCCTCGCTTTTAAAG   | CTCTATGCAAGCTTTAATCGTCTCCAATAGCTTGT      | 62045 | Db | 63066 | AATTTCTCACGAGAGGCGCTTT    | GTGTACCGAAGAATGTTTATAGCTGTGCTGCAG       | 63125 |
| Qy | 6661  | TTTAAAAATGCAACTCTGGT    | ACATCTTTATGACAGTGGTCCAAAAAACAAGATTTAA    | 6720  | Qy | 7741  | CGGGTCTGTTATACCGGCCCA     | GTGGCCAGAGTTCTGCACGACCTCTACCGGAAATGAA   | 7800  |
| Db | 62046 | TTTAAAAATGCAACTCTGGT    | ACATCTTTATGACAGTGGTCCAAAAAACAAGATTTAA    | 62105 | Db | 63126 | CGGGTCTGTTATACCGGCCCA     | GTGGCCAGAGTTCTGCACGACCTCTACCGGAAATGAA   | 63185 |
| Qy | 6721  | AAACAGCTTAACTCGCTGTCT   | GTAGTCTCTGATATAAATCTCTCAACAGAAAAA        | 6780  | Qy | 7801  | GGCCAAAGTGTGGGGGCGTGG     | CGTCTCATATGTTGAGGGCGCTATCATGGCTAT       | 7860  |
| Db | 62106 | AAACAGCTTAACTCGCTGTCT   | GTAGTCTCTGATATAAATCTCTCAACAGAAAAA        | 62165 | Db | 63186 | GGCCAAAGTGTGGGGGCGTGG     | CGTCTCATATGTTGAGGGCGCTATCATGGCTAT       | 63245 |
| Qy | 6781  | TTTTTTAACCAACTCGCCAG    | GTACTGAAACGATGGTAAGGACAGTCTGAAAGTGT      | 6840  | Qy | 7861  | TGCGACCTCGCTCTCGTCACT     | TATATACCTCGATGCCGAGGAGAACTTGAACCTTAG    | 7920  |
| Db | 62166 | TTTTTTAACCAACTCGCCAG    | GTACTGAAACGATGGTAAGGACAGTCTGAAAGTGT      | 62225 | Db | 63246 | TGCGACCTCGCTCTCGTCACT     | TATATACCTCGATGCCGAGGAGAACTTGAACCTTAG    | 63305 |
| Qy | 6841  | TATCATCTCTTACCTCCCTT    | TGAAGTGTGTTTGTGTACACCGACATGTGATAGGCA     | 6900  | Qy | 7921  | ACTCAAGCATGTGCAAACTTT     | TTTCAACCACTTTCTTCAAGAAATCTTAGCCCC       | 7980  |
| Db | 62226 | TATCATCTCTTACCTCCCTT    | TGAAGTGTGTTTGTGTACACCGACATGTGATAGGCA     | 62285 | Db | 63306 | ACTCAAGCATGTGCAAACTTT     | TTTCAACCACTTTCTTCAAGAAATCTTAGCCCC       | 63365 |
| Qy | 6901  | CCAGCTCAGGGGGAGTTCGAT   | CTGAGAGGGGTAAACAATTCGTGGCGTGGGGTACT      | 6960  | Qy | 7981  | TATGACAGCTACTGCGCAACG     | GGAAGATGTTTCTTGACAAATTTTACCATCACCGGTAC  | 8040  |
| Db | 62286 | CCAGCTCAGGGGGAGTTCGAT   | CTGAGAGGGGTAAACAATTCGTGGCGTGGGGTACT      | 62345 | Db | 63366 | TATGACAGCTACTGCGCAACG     | GGAAGATGTTTCTTGACAAATTTTACCATCACCGGTAC  | 63425 |
| Qy | 6961  | CAGTTTATAGTCTCCGAGG     | ACGCTAGAGTTCTTCTGCAAACTTACGACGCGCC       | 7020  | Qy | 8041  | GGCCGAGAGAGGGCCCTCTG      | TGTAAGTGAAGTACTAGCGGATGACGTGTCCGGGAG    | 8100  |
| Db | 62346 | CAGTTTATAGTCTCCGAGG     | ACGCTAGAGTTCTTCTGCAAACTTACGACGCGCC       | 62405 | Db | 63426 | GGCCGAGAGAGGGCCCTCTG      | TGTAAGTGAAGTACTAGCGGATGACGTGTCCGGGAG    | 63485 |
| Qy | 7021  | ATCCACAGGGGAGTACTAGT    | TTAATAGTCCCGAATTTCTATCGTCTGTAATTTGTG     | 7080  | Qy | 8101  | CAGCCTTGTCCCGATTCCTTA     | ATCCTGCGAGTATCGAACCGGGCTTCTCCCGGCACC    | 8160  |
| Db | 62406 | ATCCACAGGGGAGTACTAGT    | TTAATAGTCCCGAATTTCTATCGTCTGTAATTTGTG     | 62465 | Db | 63486 | CAGCCTTGTCCCGATTCCTTA     | ATCCTGCGAGTATCGAACCGGGCTTCTCCCGGCACC    | 63545 |
| Qy | 7081  | GAGCTCTTTAGACAGTAAC     | CAATTTGCTCCGACACTGCGGTACGATGAACCGCTTCC   | 7140  | Qy | 8161  | CCTGGTTCACCTCAGTAATGT     | CTTAGAAAAATCCAGAAATCATCTCAGCGCCCCACCCT  | 8220  |
| Db | 62466 | GAGCTCTTTAGACAGTAAC     | CAATTTGCTCCGACACTGCGGTACGATGAACCGCTTCC   | 62525 | Db | 63546 | CCTGGTTCACCTCAGTAATGT     | CTTAGAAAAATCCAGAAATCATCTCAGCGCCCCACCCT  | 63605 |
| Qy | 7141  | TTGCTGTGTTGATGACACCT    | CGAGTATCTTAGACAGGTATAGGGCCCGTTCCACTCT    | 7200  | Qy | 8221  | GAGTCAATTTGTTCATCAAA      | CACGACCCAGTCTGCTCAGTCACTCAGTCACTATTATAC | 8280  |
| Db | 62526 | TTGCTGTGTTGATGACACCT    | CGAGTATCTTAGACAGGTATAGGGCCCGTTCCACTCT    | 62585 | Db | 63606 | GAGTCAATTTGTTCATCAAA      | CACGACCCAGTCTGCTCAGTCACTCAGTCACTATTATAC | 63665 |
| Qy | 7201  | AAGGTTCCCGCTTGGTAGC     | CGGATTTGCGGGCATTTGACCGTCACTGTTCAITTCACC  | 7260  | Qy | 8281  | GCCAAACCCAGGGCGTTGTT      | CCCGGCAATGTTTATGACACGTTGGAAGCGGTGTACA   | 8340  |
| Db | 62586 | AAGGTTCCCGCTTGGTAGC     | CGGATTTGCGGGCATTTGACCGTCACTGTTCAITTCACC  | 62645 | Db | 63666 | GCCAAACCCAGGGCGTTGTT      | CCCGGCAATGTTTATGACACGTTGGAAGCGGTGTACA   | 63725 |
| Qy | 7261  | GACTATCATGCGAGTGCAGT    | AGTGGAACTTGTACATATGATTTAAAGATGCTCTT      | 7320  | Qy | 8341  | GAGCATTTACCAACAGGACAC     | AGACGCTATTTTGGCCCGCGCACTAAACCGGTTCCGATC | 8400  |
| Db | 62646 | GACTATCATGCGAGTGCAGT    | AGTGGAACTTGTACATATGATTTAAAGATGCTCTT      | 62705 | Db | 63726 | GAGCATTTACCAACAGGACAC     | AGACGCTATTTTGGCCCGCGCACTAAACCGGTTCCGATC | 63785 |
| Qy | 7321  | CCCGTAATGGGACATCATGT    | TGTTTAAACGTTTCAAGGGTGTAAAAAATCAGAAG      | 7380  | Qy | 8401  | TCGGGCGCTTGGCCGCACTAT     | CGCATGTTCCCGAGTGGCGGACAGTCTCGGACAGTGA   | 8460  |
| Db | 62706 | CCCGTAATGGGACATCATGT    | TGTTTAAACGTTTCAAGGGTGTAAAAAATCAGAAG      | 62765 | Db | 63786 | TGCGGCGCTTGGCCGCACTAT     | CGCATGTTCCCGAGTGGCGGACAGTCTCGGACAGTGA   | 63845 |
| Qy | 7381  | GTCCCTGCGTGAACACTTTT    | TCGGACTCTCCGACCGCTCAGGATTAAGCTGTACACGC   | 7440  | Qy | 8461  | AGGGTCCGCGGCGTCTGAT       | TGCGGAAACACGACATGCTTCGAGCCGCGGTACCGCC   | 8520  |
| Db | 62766 | GTCCCTGCGTGAACACTTTT    | TCGGACTCTCCGACCGCTCAGGATTAAGCTGTACACGC   | 62825 | Db | 63846 | AGGGTCCGCGGCGTCTGAT       | TGCGGAAACACGACATGCTTCGAGCCGCGGTACCGCC   | 63905 |
|    |       |                         |  |       | Qy | 8521  | CAATGTCCCGCCACGAGGAT      | GGAACGGTAGCACAGTTGGGAAACGCTCCCGTAAAAA   | 8580  |

Db 63906 CAATGTCCTCCGCGCAAGAGAGTGAAGACGGTAGCACAGTTGGGAAACGCTCCCGTAAATAA 63965  
Qy 8381 CGTCCACATCGGAGCGCGGTATACGCTCCACTGGTTATATACCAATATAGACTTAAC 8640  
Db 63966 CGTCCACATCGGAGCGCGGTATACGCTCCACTGGTTATATACCAATATAGACTTAAC 64025  
Qy 8641 GTCCCGCTCAGGGTTCGGCCAGAGTCGGCCGATATCGCCAACTCCAGAGTCCCGCAT 8700  
Db 64026 GTCCCGCTCAGGGTTCGGCCAGAGTCGGCCGATATCGCCAACTCCAGAGTCCCGCAT 64085  
Qy 8701 GCGCGCGGCTCTCCGCGCTTCGCGCAAAACCGCGCAACGCTCCCGCTTAGAGAAAGCA 8760  
Db 64086 GCGCGCGGCTCTCCGCGCTTCGCGCAAAACCGCGCAACGCTCCCGCTTAGAGAAAGCA 64145  
Qy 8761 GCGCGCGGAGCGTGGCAGACAAAGAGTGAAGGCGGACGTTCCGGCGCGCGCAACAGT 8820  
Db 64146 GCGCGCGGAGCGTGGCAGACAAAGAGTGAAGGCGGACGTTCCGGCGCGCGCAACAGT 64205  
Qy 8821 AATCCACCTTTCCCGGACCGTCCGGATGCGGTTCCGAGCAGGCGCTTATTCGATT 8880  
Db 64206 AATCCACCTTTCCCGGACCGTCCGGATGCGGTTCCGAGCAGGCGCTTATTCGATT 64265  
Qy 8881 AATCGAAAGTCCACGATGTAAACCGCAACGATCTGGACCCAAACGACGACGACAT 8940  
Db 64266 AATCGAAAGTCCACGATGTAAACCGCAACGATCTGGACCCAAACGACGACGACAT 64325  
Qy 8941 GCTAGCGGCTATCTACAGACGCTGTATGACTACAGTCCCGCGGCGATTCGATTCGCC 9000  
Db 64326 GCTAGCGGCTATCTACAGACGCTGTATGACTACAGTCCCGCGGCGATTCGATTCGCC 64385  
Qy 9001 CTCAGCAACTCGGCAATGAGAGATATTTCCAGAGGTAGTCCGCCATCTAGCGGCA 9060  
Db 64386 CTCAGCAACTCGGCAATGAGAGATATTTCCAGAGGTAGTCCGCCATCTAGCGGCA 64445  
Qy 9061 CGGATCGCCTTGA 9073  
Db 64446 CGGATCGCCTTGA 64458

RESULT 2  
AAC64754  
ID AAC64754 standard; DNA; 133719 BP.  
XX  
AC AAC64754;  
XX  
DT 28-FEB-2001 (first entry)  
XX  
DE Macaca mulatta rhadinovirus 17577 (RRV) genome sequence SEQ ID NO:1.  
XX  
KW Macaca mulatta rhadinovirus 17577; RRV; rhesus macaque rhadinovirus;  
KW genome; Kaposi's sarcoma-associated herpesvirus; KSHV; interleukin 6;  
KW IL-6; macrophage inflammatory protein; MIP; diagnosis; vaccine;  
KW cytostatic; anti-HIV; gene therapy; infection; Kaposi's sarcoma;  
KW lymphoproliferative disorder; B-cell hyperplasia; lymphadenopathy;  
KW splenomegaly; hypergammaglobulinemia; autoimmune haemolytic anaemia;  
KW ds.  
XX  
OS Macaca mulatta rhadinovirus 17577.  
XX  
PN W0200028040-A2.  
XX  
PD 18-MAY-2000.  
XX  
PF 05-NOV-1999; 99WO-US026260.  
XX  
PR 06-NOV-1998; 98US-0107507P.  
PR 20-NOV-1998; 98US-0109409P.  
XX  
PA (UOYOR-) UNIV OREGON HEALTH SCI.  
XX  
PI Wong SW, Axthelm MK, Searles RP;  
XX

WPI; 2000-376552/32.  
New rhesus rhadino virus for producing non-human primate model useful for testing potential treatments and efficacy of the candidate vaccine for conditions associated with RRV infection.  
Claim 2; Page 83-122; 141pp; English.  
The present invention describes a novel rhesus macaque rhadinovirus called macaca mulatta rhadinovirus 17577 (RRV). AAC64754 represents the RRV genome sequence, and AAB53123 to AAB53204 represent the proteins encoded by the genome sequence. The present invention also specifically claims the individual open reading frame (ORF) nucleotide sequences from the genome which encode the individual proteins, but these sequences are not given. A non-human animal infected with RRV can be used for testing the efficacy of drug in the treatment of condition associated with infection with RRV such as Kaposi's sarcoma, lymphoproliferative disorders, B-cell hyperplasia, lymphadenopathy, splenomegaly, hypergammaglobulinemia or autoimmune haemolytic anaemia, by administering the drug to an immuno-compromised non-human primate preferably Rhesus macaque monkey obtained by as a result of infection by Simian Immunodeficiency Virus (SIV). RRV is useful for producing non-human primate model for testing potential treatments for conditions associated with RRV infection. It is also useful for testing the efficacy of the candidate vaccine against RRV infection or conditions associated with its infection by administering the vaccine to the subject capable of infection with RRV, inoculating the subject with RRV and observing the effect of vaccine. AAC64755 to AAC64765 and AAB53205 to AAB53213 represent sequence used in the exemplification of the present invention  
Sequence 133719 BP; 32746 A; 35648 C; 34521 G; 30804 T; 0 U; 0 Other;  
Query Match 100.0%; Score 9073; DB 3; Length 133719;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 9073; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGAGAGTTCGTCGATGGACCAACACGTCGAAACCAATCCGGGTTTCATTTGAAC 60  
Db 60966 ATGGAGAGTTCGTCGATGGACCAACACGTCGAAACCAATCCGGGTTTCATTTGAAC 61025  
Qy 61 ATGAGCTCCGATGCCAAAGTCAGGGGTGTCGTGATCAGTCAGTCGCTGTCTCAATATA 120  
Db 61026 ATGAGCTCCGATGCCAAAGTCAGGGGTGTCGTGATCAGTCAGTCGCTGTCTCAATATA 61085  
Qy 121 ACTACAGCCACCGGAAATGGGTGGTACGACCTGCGCTTCGATCCGGGTGAAGATCC 180  
Db 61086 ACTACAGCCACCGGAAATGGGTGGTACGACCTGCGCTTCGATCCGGGTGAAGATCC 61145  
Qy 181 GGGCGGTTCTTGGCGTTTACCGTTTATCTAATTACGGAACTGCTGCTGGGAAAGT 240  
Db 61146 GGGCGGTTCTTGGCGTTTACCGTTTATCTAATTACGGAACTGCTGCTGGGAAAGT 61205  
Qy 241 ACCAGCATATCGGCGCTGTACCAAAATTTAACTGCTGTATCAGGGCGGACCAACATA 300  
Db 61206 ACCAGCATATCGGCGCTGTACCAAAATTTAACTGCTGTATCAGGGCGGACCAACATA 61265  
Qy 301 GCGGCACAGAACCTATCGCTCGCTTAAAGAGCTTCTGTCCACGATCTTTCAGGCGTTT 360  
Db 61266 GCGGCACAGAACCTATCGCTCGCTTAAAGAGCTTCTGTCCACGATCTTTCAGGCGTTT 61325  
Qy 361 GCGTTTAAAGAGCGGACATCAATATAGCCGTAGAAAGAGCTATCAGACCGGAGCCGTA 420  
Db 61326 GCGTTTAAAGAGCGGACATCAATATAGCCGTAGAAAGAGCTATCAGACCGGAGCCGTA 61385  
Qy 421 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGGCTGCTTAAGAGCTTATAGTGGACAT 480  
Db 61386 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGGCTGCTTAAGAGCTTATAGTGGACAT 61445  
Qy 481 ATGAAGAGGTTATGGCGAAACCAATGGCGATGTACGGGACTATATCAACGGGAAT 540  
Db 61446 ATGAAGAGGTTATGGCGAAACCAATGGCGATGTACGGGACTATATCAACGGGAAT 61505  
Qy 541 TTTGAAACCCCTCTCGAGAAATGACCGGACCGTGTATTATGGAGCTTCCATATATTATTAATC 600



Db 61506 TTTGAAACCTCTCGAGATGACCGGACCGTGTATTGAGACTTCCAAATATTATGTAATC 61565  
Qy 601 GACGAGCGGAAACCTCTGCTCTTACATACCTACACCGTGTGTTCTTTACTGGTTC 660  
Db 61566 GACGAGCGGAAACCTCTGCTCTTACATACCTACACCGTGTGTTCTTTACTGGTTC 61625  
Qy 661 CTAAACAGCTGCTGCTAAATACCTCTTACCGGACGCGGCGGTTCGGTGCATGATGC 720  
Db 61626 CTAAACAGCTGCTGCTAAATACCTCTTACCGGACGCGGCGGTTCGGTGCATGATGC 61685  
Qy 721 GTCGGTTCGCAACGCGAGAAACCGCTTCCAGTCAACTTACCAACACGCGGACGCAAAAG 780  
Db 61686 GTCGGTTCGCAACGCGAGAAACCGCTTCCAGTCAACTTACCAACACGCGGACGCAAAAG 61745  
Qy 781 ACGAGATATGCTGCTGCGAGAACATCTTAACATTCATGATAGGAGAGAGTCTGATCT 840  
Db 61746 ACGAGATATGCTGCTGCGAGAACATCTTAACATTCATGATAGGAGAGAGTCTGATCT 61805  
Qy 841 GAGTACGTPACATTTGGAGAGAACTGGCGCTGTTTATAAACAAGCGCTGCACCGAT 900  
Db 61806 GAGTACGTPACATTTGGAGAGAACTGGCGCTGTTTATAAACAAGCGCTGCACCGAT 61865  
Qy 901 CTGCAGTTCGACACCTCTTAATAATTTAGAGTATAATCTTCCCAATCTCTGACGAGTC 960  
Db 61866 CTGCAGTTCGACACCTCTTAATAATTTAGAGTATAATCTTCCCAATCTCTGACGAGTC 61925  
Qy 961 ATGAGTACGTPACGATTTGCTGCTTAAAGTAAGATTAAGATCTTTAGAAATAC 1020  
Db 61926 ATGAGTACGTPACGATTTGCTGCTTAAAGTAAGATTAAGATCTTTAGAAATAC 61985  
Qy 1021 ATTGGCTGACCCGACTCTTTGTACATAGCGAGTAAAGCGTATCTTAACAACCTA 1080  
Db 61986 ATTGGCTGACCCGACTCTTTGTACATAGCGAGTAAAGCGTATCTTAACAACCTA 62045  
Qy 1081 CACACATGCTPAACGCTAGGCGGATACGAGGACACGAAGCTCTTTACCTGTCCCGTG 1140  
Db 62046 CACACATGCTPAACGCTAGGCGGATACGAGGACACGAAGCTCTTTACCTGTCCCGTG 62105  
Qy 1141 GTGTGCGAGGTGTTTGTGAAGCGTGTGAGGAATACAGGCGGCGTCAACCTCAACAC 1200  
Db 62106 GTGTGCGAGGTGTTTGTGAAGCGTGTGAGGAATACAGGCGGCGTCAACCTCAACAC 62165  
Qy 1201 CTCACGCTGACGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATATTGCGAGTTT 1260  
Db 62166 CTCACGCTGACGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATATTGCGAGTTT 62225  
Qy 1261 GTGACACGAGCATGTCCATAGTCGCCACGGAATCCACTGAAGCTTTCGACCCAGTTACC 1320  
Db 62226 GTGACACGAGCATGTCCATAGTCGCCACGGAATCCACTGAAGCTTTCGACCCAGTTACC 62285  
Qy 1321 TTTATACCAAGTTTGTAAACACGACCGTCTTTAAATGGAAGAAACAAAAATGC 1380  
Db 62286 TTTATACCAAGTTTGTAAACACGACCGTCTTTAAATGGAAGAAACAAAAATGC 62345  
Qy 1381 ATATGTGGTTTCAGGTTAGTACTTTTGTAGTTCGAAAGAAATCTAGACAGGAACTCTTC 1440  
Db 62346 ATATGTGGTTTCAGGTTAGTACTTTTGTAGTTCGAAAGAAATCTAGACAGGAACTCTTC 62405  
Qy 1441 GTGAGACGCAATTCGAAAGCGCTGCGGAATACGTATACGGTTTCTTAAACACGTTACTG 1500  
Db 62406 GTGAGACGCAATTCGAAAGCGCTGCGGAATACGTATACGGTTTCTTAAACACGTTACTG 62465  
Qy 1501 TACAACGCGCATGTACTGTTTACGCGTACGCGTGTAGTGTACACAGAGAAATATCTG 1560  
Db 62466 TACAACGCGCATGTACTGTTTACGCGTACGCGTGTAGTGTACACAGAGAAATATCTG 62525  
Qy 1561 CAAGACCTAAATTTGACCCCTCCCGCGCTCTGGCAACCGGCGGTGTAGACCTTCAA 1620  
Db 62526 CAAGACCTAAATTTGACCCCTCCCGCGCTCTGGCAACCGGCGGTGTAGACCTTCAA 62585  
Qy 1621 ACGGTTGCTGAAGAGTTTAAACCTTGAAGACGACATCTTTACACCGTGTGAGTCCCGG 1680

Db 62586 ACGGTTGCTGAAGAGTTAAACCTGGAAGACGACATCTTCTACACGTTGTCAGTCCCGG 62645  
Qy 1681 CCAACCCCGGTTATCACTCTCCCTCCAGGTTTGGTGCACAGTACTGCGCCCTAAAGAC 1740  
Db 62646 CCAACCCCGGTTATCACTCTCCCTCCAGGTTTGGTGCACAGTACTGCGCCCTAAAGAC 62705  
Qy 1741 GTGTTGCGCTCCAGAAATAAGGTGCGTCTGTTTGGCGGAGTGTGAGAGAA 1800  
Db 62706 GTGTTGCGCTCCAGAAATAAGGTGCGTCTGTTTGGCGGAGTGTGAGAGAA 62765  
Qy 1801 ACGTTTTCGCGTTTACCGTTAACATGTCGTAAGGACGAGTTGACTTTGTCTCCCT 1860  
Db 62766 ACGTTTTCGCGTTTACCGTTAACATGTCGTAAGGACGAGTTGACTTTGTCTCCCT 62825  
Qy 1861 TCAGAAAGCTCTCAACCGGCTGTTGGCGTTGCATCGACGCTTGAATCGTATATAAATTAA 1920  
Db 62826 TCAGAAAGCTCTCAACCGGCTGTTGGCGTTGCATCGACGCTTGAATCGTATATAAATTAA 62885  
Qy 1921 GGTTACACGTTTTCACCGTAGCGTCTGTCAGGTTTCCACTCAGTATGATAC 1980  
Db 62886 GGTTACACGTTTTCACCGTAGCGTCTGTCAGGTTTCCACTCAGTATGATAC 62945  
Qy 1981 CTCAGAAAGATGCGCTCCCTCCCTGCTGTCAGGACTCTAGCGGTTTATCGCGTCCCTA 2040  
Db 62946 CTCAGAAAGATGCGCTCCCTCCCTGCTGTCAGGACTCTAGCGGTTTATCGCGTCCCTA 63005  
Qy 2041 GAGAAATAACATTAACCAATTTGACCGAAGCATGAGAGGACGAGGCTTTTCAAGTGTGC 2100  
Db 63006 GAGAAATAACATTAACCAATTTGACCGAAGCATGAGAGGACGAGGCTTTTCAAGTGTGC 63065  
Qy 2101 TGTGCGGCGACTATGCGGTGAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGGA 2160  
Db 63066 TGTGCGGCGACTATGCGGTGAGCTCAAAATTTAGCCATGACCATCGTAAAGGACACAGGGA 63125  
Qy 2161 ATGTGCTGAGGAGTACCGTGTAGTATTTGGTCTCCACAGAAAGTTCAGACAGCCAC 2220  
Db 63126 ATGTGCTGAGGAGTACCGTGTAGTATTTGGTCTCCACAGAAAGTTCAGACAGCCAC 63185  
Qy 2221 GTGTATGTAGCAATATCAAGGCTGTAACTCAAAATTTATTTGGTCAATGACGACGACCC 2280  
Db 63186 GTGTATGTAGCAATATCAAGGCTGTAACTCAAAATTTATTTGGTCAATGACGACGACCC 63245  
Qy 2281 CTTTAAACCTCTCTCAGAGAACAGTTCGATAACCTCGGCCACGATATAGTCCGCGC 2340  
Db 63246 CTTTAAACCTCTCTCAGAGAACAGTTCGATAACCTCGGCCACGATATAGTCCGCGC 63305  
Qy 2341 CTCACAAACCCCAACCAACCTCTACTATAATAAATAAATAAATAAATAAATAAATAAATAA 2400  
Db 63306 CTCACAAACCCCAACCAACCTCTACTATAATAAATAAATAAATAAATAAATAAATAAATAA 63365  
Qy 2401 ATTGTGTTTTTATCAGTCCAAACGAGGACAGGATAGTGTCTATTTTCCACACCGGG 2460  
Db 63366 ATTGTGTTTTTATCAGTCCAAACGAGGACAGGATAGTGTCTATTTTCCACACCGGG 63425  
Qy 2461 GGGTGGCCGCTATAGTTTACGACCAAGAAAGCGCCGCGCTCTGGGTTAGAAAGTAGGCC 2520  
Db 63426 GGGTGGCCGCTATAGTTTACGACCAAGAAAGCGCCGCGCTCTGGGTTAGAAAGTAGGCC 63485  
Qy 2521 CGCCCCCGGTGTGAGGTCATACGTCTTTAGGCGCTGGGTTGTACCAACATAAATCA 2580  
Db 63486 CGCCCCCGGTGTGAGGTCATACGTCTTTAGGCGCTGGGTTGTACCAACATAAATCA 63545  
Qy 2581 CGTGTGTATCAGCTGTGAGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2640  
Db 63546 CGTGTGTATCAGCTGTGAGTCTTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 63605  
Qy 2641 CGCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 2700  
Db 63606 CGCGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 63665  
Qy 2701 CATCAGTGGCGCGCTGCTCATCGGCTGTGATTTGCTGCTCCATGCTCTCAACAGAGCTC 2760  
Db 63666 CATCAGTGGCGCGCTGCTCATCGGCTGTGATTTGCTGCTCCATGCTCTCAACAGAGCTC 63725

QY 2761 GGTGCTGCACTTCGTCGTCATCTGAAATAATAATGCTCTCTGTCTCTGCTGCTGATCCAG 2820  
Db 63726 GGTGCTGCACTTCGTCGTCATCTGAAATAATAATGCTCTCTGTCTCTGCTGATCCAG 63785  
QY 2821 CAACAACAGGTTCTTCTTCATCGGAAGACCATCCGCCCTCTGCGGTGATCTAGATATAT 2880  
Db 63786 CAACAACAGGTTCTTCTTCATCGGAAGACCATCCGCCCTCTGCGGTGATCTAGATATAT 63845  
QY 2881 CACCTGACCAAGCGGCGAGGCGCGTGGGGAAGCGCTGAGCTTCTCCGCCACTTTCAT 2940  
Db 63846 CACCTGACCAAGCGGCGAGGCGCGTGGGGAAGCGCTGAGCTTCTCCGCCACTTTCAT 63905  
QY 2941 CAGTGTCTCTGCTGATCTATAAAAAATCTCTTCGCTGCTGCTGATTCAGATGACTGGG 3000  
Db 63906 CAGTGTCTCTGCTGATCTATAAAAAATCTCTTCGCTGCTGCTGATTCAGATGACTGGG 63965  
QY 3001 GTGCGCGGTGTTGGCCAGTAACACCGCAACGGCGCGGACCAACCAACCGCGTCACTG 3060  
Db 63966 GTGCGCGGTGTTGGCCAGTAACACCGCAACGGCGCGGACCAACCAACCGCGTCACTG 64025  
QY 3061 GCGCGCGCGCTCGGTGCTGTGGTGGCACCTGTGCAAGTAATGACGCTAGGTAGCGCGC 3120  
Db 64026 GCGCGCGCGCTCGGTGCTGTGGTGGCACCTGTGCAAGTAATGACGCTAGGTAGCGCGC 64085  
QY 3121 CGCTCTGGTTAAGGACTCACCGTCCCGTGGTGGATCTATTCTAGTGGGAGGAGA 3180  
Db 64086 CGCTCTGGTTAAGGACTCACCGTCCCGTGGTGGATCTATTCTAGTGGGAGGAGA 64145  
QY 3181 CCGCGCGGTTCGTTCTTTCATCCACATCACTAGTCTGCTGCTCCGAATGAGAGTCGG 3240  
Db 64146 CCGCGCGGTTCGTTCTTTCATCCACATCACTAGTCTGCTGCTCCGAATGAGAGTCGG 64205  
QY 3241 TTTCTGCTCGGTTTGTGAGTAGTGGGAGGGGATCTTGGTCAAAAACATCTCAGTGTG 3300  
Db 64206 TTTCTGCTCGGTTTGTGAGTAGTGGGAGGGGATCTTGGTCAAAAACATCTCAGTGTG 64265  
QY 3301 CAGTCACTGCAATCACTGGGCGCTCCGTACCGGGCGGTACCGCAACCGCGTACGGG 3360  
Db 64266 CAGTCACTGCAATCACTGGGCGCTCCGTACCGGGCGGTACCGCAACCGCGTACGGG 64325  
QY 3361 ACCCGCGAACAAAAACGTCCTGTTCTTTTCTAGGTGCTCCCGGGAATCGGAGCA 3420  
Db 64326 ACCCGCGAACAAAAACGTCCTGTTCTTTTCTAGGTGCTCCCGGGAATCGGAGCA 64385  
QY 3421 TCTTGGAGTGGCGCGGGGCTTTCAGGGTGGATGAAGCAATAGCCATGGCGGAA 3480  
Db 64386 TCTTGGAGTGGCGCGGGGCTTTCAGGGTGGATGAAGCAATAGCCATGGCGGAA 64445  
QY 3481 CGTTCACTGTAAACGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGGCAC 3540  
Db 64446 CGTTCACTGTAAACGCACTGCTATCCCGATCAGTTAATATTCAGTCAACGGCAC 64505  
QY 3541 CTTCCGATTCGAACTAGATAGTCACTCGAAGTTTAAATGGTTGACGCTAGAACTG 3600  
Db 64506 CTTCCGATTCGAACTAGATAGTCACTCGAAGTTTAAATGGTTGACGCTAGAACTG 64565  
QY 3601 CGGCCAGGAGAACCGGTTGCGCGTGGCGGTAGCCAGGAGACGGATGGCGGCTT 3660  
Db 64566 CGGCCAGGAGAACCGGTTGCGCGTGGCGGTAGCCAGGAGACGGATGGCGGCTT 64625  
QY 3661 GAGGACGAGATGCTCTGTCGTTTATCAGCGTTCGCTTCTATAGCTTCTTCCCA 3720  
Db 64626 GAGGACGAGATGCTCTGTCGTTTATCAGCGTTCGCTTCTATAGCTTCTTCCCA 64685  
QY 3721 CAGCATAAACCGACTCTGAGCTTGGCAGAAAGCGACTTATATGCTAGCTGTGAA 3780  
Db 64686 CAGCATAAACCGACTCTGAGCTTGGCAGAAAGCGACTTATATGCTAGCTGTGAA 64745  
QY 3781 CCAATCCAGCGAGGTTGGCGTGTGACCCCGCTTCCCTCTTTCCACCGTCAAGATG 3840  
Db 64746 CCAATCCAGCGAGGTTGGCGTGTGACCCCGCTTCCCTCTTTCCACCGTCAAGATG 64805

QY 3841 GTTTAAAGCAGAACTCCCGTTTTGCCCAACAGTCTAAGCACCCGTGAGAGAGCGGT 3900  
Db 64806 GTTTAAAGCAGAACTCCCGTTTTGCCCAACAGTCTAAGCACCCGTGAGAGAGCGGT 64865  
QY 3901 GAACCCAGGTACAGTATTCGCTATCTCTTTTAAAAATATTTTGGAGCTTCGAGGTATCT 3960  
Db 64866 GAACCCAGGTACAGTATTCGCTATCTCTTTTAAAAATATTTTGGAGCTTCGAGGTATCT 64925  
QY 3961 GTAGTCCGGAGCAACGCTAAGAGCTAGTCCGGTGTGACCCGCGATGGTAAAGATCTG 4020  
Db 64926 GTAGTCCGGAGCAACGCTAAGAGCTAGTCCGGTGTGACCCGCGATGGTAAAGATCTG 64985  
QY 4021 GCCAAGATCAGACACTTAAATATCTCCGGTTCGCAACAAATAGACACCACATATCT 4080  
Db 64986 GCCAAGATCAGACACTTAAATATCTCCGGTTCGCAACAAATAGACACCACATATCT 65045  
QY 4081 ATCTATCCGTGGTAATAATCAGCGTTACGTTGCTCATATCCATTAACCGCTTCAGAGAGC 4140  
Db 65046 ATCTATCCGTGGTAATAATCAGCGTTACGTTGCTCATATCCATTAACCGCTTCAGAGAGC 65105  
QY 4141 CGGAGCTTCTGTTTTTAAAAATCGAAGATTTAAAAATTCAGCCATGCTCGCTAAG 4200  
Db 65106 CGGAGCTTCTGTTTTTAAAAATCGAAGATTTAAAAATTCAGCCATGCTCGCTAAG 65165  
QY 4201 TAGCAAGCTTGTGTGGAGGTTCTTCTTAACACTTCTGTGTGACATTTTACTCCAAACTAT 4260  
Db 65166 TAGCAAGCTTGTGTGGAGGTTCTTCTTAACACTTCTGTGTGACATTTTACTCCAAACTAT 65225  
QY 4261 TGTTTAAGCAACCTTCCATAAATCCCGCGTCTGTGTAACTTTACGTATACGTGAA 4320  
Db 65226 TGTTTAAGCAACCTTCCATAAATCCCGCGTCTGTGTAACTTTACGTATACGTGAA 65285  
QY 4321 ACTTTTGTACTTCTGCTGAGAAATCCACAGAGGTGGCGCGTATAAGCAGTTCGCTT 4380  
Db 65286 ACTTTTGTACTTCTGCTGAGAAATCCACAGAGGTGGCGCGTATAAGCAGTTCGCTT 65345  
QY 4381 TCACTACCGTTTATGTGAAGTCTTAAACTCGGCTAAAAAATCATCAAGACTATATTC 4440  
Db 65346 TCACTACCGTTTATGTGAAGTCTTAAACTCGGCTAAAAAATCATCAAGACTATATTC 65405  
QY 4441 ATATAATTCACAAATTCAGCTTCTCTCGGGACACTGTAATTTAACTTGTACAAATC 4500  
Db 65406 ATATAATTCACAAATTCAGCTTCTCTCGGGACACTGTAATTTAACTTGTACAAATC 65465  
QY 4501 CTTTAAAAAAGCGCTTAAGTGAAATCCATTTTACATATTTCTGTGGTAGAGCA 4560  
Db 65466 CTTTAAAAAAGCGCTTAAGTGAAATCCATTTTACATATTTCTGTGGTAGAGCA 65525  
QY 4561 GCTTTATATCGCAATGTGCTGTACACAGCCCTGGACATAAGCTAAATTTGTAAATGA 4620  
Db 65526 GCTTTATATCGCAATGTGCTGTACACAGCCCTGGACATAAGCTAAATTTGTAAATGA 65585  
QY 4621 AAAGAAGCAATCTCAACCGCTCCGGAATATAAAGTCACTAAGTGGTATTTTACAA 4680  
Db 65586 AAAGAAGCAATCTCAACCGCTCCGGAATATAAAGTCACTAAGTGGTATTTTACAA 65645  
QY 4681 CATGGATTTACTACAGCCCGGTACAACTGTAAATTTAAAAAATCCACATGTAAAT 4740  
Db 65646 CATGGATTTACTACAGCCCGGTACAACTGTAAATTTAAAAAATCCACATGTAAAT 65705  
QY 4741 AAACTTAAAGTATACATCTTCTCATGTTTTTATACATAATATGAACGCTAATGCTACT 4800  
Db 65706 AAACTTAAAGTATACATCTTCTCATGTTTTTATACATAATATGAACGCTAATGCTACT 65765  
QY 4801 TATATAATGTATAGAAATTAATGCCAATACAGTCACTGTAAGTATATTTTAAAGCAAT 4860  
Db 65766 TATATAATGTATAGAAATTAATGCCAATACAGTCACTGTAAGTATATTTTAAAGCAAT 65825  
QY 4861 GCAACATCAAAATATATACAAAAATGACGCTTCGGATATCCGCCAGCTGTAAAC 4920  
Db 65826 GCAACATCAAAATATATACAAAAATGACGCTTCGGATATCCGCCAGCTGTAAAC 65885  
QY 4921 CCGAAAAATACAGAAATGACTACACAAACACACCTGAAACCAATTTTATCTTCAACATATG 4980



QY 7141 TTGTCGTGTTGATAGACACCTGAGATATCTTAGACAGGTATAGGGCCCGTTCCACTCT 7200  
DB 68106 TTGTCGTGTTGATAGACACCTGAGATATCTTAGACAGGTATAGGGCCCGTTCCACTCT 68165  
QY 7201 AAGGTTTCGGTTCGTTAGAGCGGATTCGGGGCATTCAGTCAGTGTGTTCAATTCACC 7260  
DB 68166 AAGGTTTCGGTTCGTTAGAGCGGATTCGGGGCATTCAGTCAGTGTGTTCAATTCACC 68225  
QY 7261 GACTATCATCGCGATGCGGATAGTGAACCTTGTACATATGATTTAAAGATAGTCTTT 7320  
DB 68226 GACTATCATCGCGATGCGGATAGTGAACCTTGTACATATGATTTAAAGATAGTCTTT 68285  
QY 7321 CCGGTAATGCGACTCATGTCGTGTTTAAAGTTTCAAGGTTTAAAGATAGTCTTT 7380  
DB 68286 CCGGTAATGCGACTCATGTCGTGTTTAAAGTTTCAAGGTTTAAAGATAGTCTTT 68345  
QY 7381 GTCTCGGTGGAACAATTTTGGGACTCTCCGACCGCTCAGGGATAGCTGTACACGCG 7440  
DB 68346 GTCTCGGTGGAACAATTTTGGGACTCTCCGACCGCTCAGGGATAGCTGTACACGCG 68405  
QY 7441 GGTGTTCTTTTAAACGAAGNATCGTCAAAAAATGACCCACGTGAATTAAGCT 7500  
DB 68406 GGTGTTCTTTTAAACGAAGNATCGTCAAAAAATGACCCACGTGAATTAAGCT 68465  
QY 7501 CACAGAAGTCTCTCGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCATCTCCGG 7560  
DB 68466 CACAGAAGTCTCTCGATGATTTGTAATGGAATGTCCTCCCTGGGTCCCATCTCCGG 68525  
QY 7561 GCTCATCGCGGATCTGAATCTCTTTAACTGTTCTGCCTCTACCGGGTTCCTCCGGTTAA 7620  
DB 68526 GCTCATCGCGGATCTGAATCTCTTTAACTGTTCTGCCTCTACCGGGTTCCTCCGGTTAA 68585  
QY 7621 AACCCGGGGCGGCACCTGTAAGTTCGGTCGGCAGAAATGCGCCAGGGCATCTGTAG 7680  
DB 68586 AACCCGGGGCGGCACCTGTAAGTTCGGTCGGCAGAAATGCGCCAGGGCATCTGTAG 68645  
QY 7681 AATTCTCACCGAGAGGGCCCTTGTGTACCGAAAGATGTTTCATAGCGTCTGCTGCAG 7740  
DB 68646 AATTCTCACCGAGAGGGCCCTTGTGTACCGAAAGATGTTTCATAGCGTCTGCTGCAG 68705  
QY 7741 CCGGGTGTGTTACCGCCCGAGTTCGGCAGAGTCTTCGACAGCTCTACGGGAAATGAA 7800  
DB 68706 CCGGGTGTGTTACCGCCCGAGTTCGGCAGAGTCTTCGACAGCTCTACGGGAAATGAA 68765  
QY 7801 GSCCAAGTGTGGGGGTGGCGTCGTCTCATATGTTGAGGGGCGCTATCATGGCTAT 7860  
DB 68766 GSCCAAGTGTGGGGGTGGCGTCGTCTCATATGTTGAGGGGCGCTATCATGGCTAT 68825  
QY 7861 TSCCGACTCGTCTCGTCACTTATATAACCTTGGATCCGAGGGAATCTGGAATCTAG 7920  
DB 68826 TSCCGACTCGTCTCGTCACTTATATAACCTTGGATCCGAGGGAATCTGGAATCTAG 68885  
QY 7921 ACTCAAGCACTGTGCAACTGTTTTCACCCCATCTTCTTCAAGAAATCTTAGCCCC 7980  
DB 68886 ACTCAAGCACTGTGCAACTGTTTTCACCCCATCTTCTTCAAGAAATCTTAGCCCC 68945  
QY 7981 TATCAGCTACTGGCCCAACCGGAAGATGTTTCTGACAAATTTATTTTACCATCAGCGTAC 8040  
DB 68946 TATCAGCTACTGGCCCAACCGGAAGATGTTTCTGACAAATTTATTTTACCATCAGCGTAC 69005  
QY 8041 GSCCGAAGAGCGGCCCTGTCGTGACTGGAAAGTACTAGCGGATGACGTGTCGGGAAG 8100  
DB 69006 GSCCGAAGAGCGGCCCTGTCGTGACTGGAAAGTACTAGCGGATGACGTGTCGGGAAG 69065  
QY 8101 CAGCCTTGTCCCGATTCTTAACTCTGACGATGCGAATCCGGGGCTTCTCCCGGCACC 8160  
DB 69066 CAGCCTTGTCCCGATTCTTAACTCTGACGATGCGAATCCGGGGCTTCTCCCGGCACC 69125  
QY 8161 CTGTGTTGACCTCAGTAAATCTTAGAAAAATCAGAAATATCTCAGCGGCCCAACCTT 8220  
DB 69126 CTGTGTTGACCTCAGTAAATCTTAGAAAAATCAGAAATATCTCAGCGGCCCAACCTT 69185

QY 8221 GAGTCAATTTGTCTATCAGAAAAACGACCCAGTCTGCCTCAGTCAGTCAGCAATTATTAC 8280  
DB 69186 GAGTCAATTTGTCTATCAGAAAAACGACCCAGTCTGCCTCAGTCAGTCAGTCAGCAATTATTAC 69245  
QY 8281 GCCAACCCAGGGGTTGTTCCCGGCCAATGTTTATGGAACAGTGTGAAAGCGGTGTCA 8340  
DB 69246 GCCAACCCAGGGGTTGTTCCCGGCCAATGTTTATGGAACAGTGTGAAAGCGGTGTCA 69305  
QY 8341 GAGCAATCACACAGGCAAGACGCTATTTTGGCGCGGCACTAACCGGTTTCGACATC 8400  
DB 69306 GAGCAATCACACAGGCAAGACGCTATTTTGGCGCGGCACTAACCGGTTTCGACATC 69365  
QY 8401 TGGCGGCCCTGGGCCCATATGCTATGTTTCCCGAGTTGCGGCACTGTCGCGCAGGTGA 8460  
DB 69366 TGGCGGCCCTGGGCCCATATGCTATGTTTCCCGAGTTGCGGCACTGTCGCGCAGGTGA 69425  
QY 8461 AGGTCGCGGGCGTCTGATTTGCGGAAACAGCATGCTTCCGACGCCGCGGTTACCGCC 8520  
DB 69426 AGGTCGCGGGCGTCTGATTTGCGGAAACAGCATGCTTCCGACGCCGCGGTTACCGCC 69485  
QY 8521 CATGTCGCCGCCAAGAGGATGGAACGGTGTGACACAGTTGGGAAACGCTCCCGTAAAAA 8580  
DB 69486 CAATGTCCCGGCCAAGAGGATGGAACGGTGTGACACAGTTGGGAAACGCTCCCGTAAAAA 69545  
QY 8581 CGTCCCATCGAGGCGCGGTATACGCTCCACTGTTTAAATATATACCAATATAGACTTAAC 8640  
DB 69546 CGTCCCATCGAGGCGCGGTATACGCTCCACTGTTTAAATATATACCAATATAGACTTAAC 69605  
QY 8641 GTCCCGTCTAGGTCGCGCCAGAGTTCGGCCGATATGCGCAACACTCCAGAGTCCCGCAT 8700  
DB 69606 GTCCCGTCTAGGTCGCGCCAGAGTTCGGCCGATATGCGCAACACTCCAGAGTCCCGCAT 69665  
QY 8701 GCGCGCGGCTCTCCGCCCTTCCGCAAAACCGCCGCAACGGTCCCGCTAAAGAGAAAGCA 8760  
DB 69666 GCGCGCGGCTCTCCGCCCTTCCGCAAAACCGCCGCAACGGTCCCGCTAAAGAGAAAGCA 69725  
QY 8761 GCCACGCGAGGACGTGGCAGACAAAGACTGAAGGGCGAGCTTCGGGGCGCCGACAGT 8820  
DB 69726 GCCACGCGAGGACGTGGCAGACAAAGACTGAAGGGCGAGCTTCGGGGCGCCGACAGT 69785  
QY 8821 ABAACACCTTTCCCGGACCGTCCGGATGCGCGTTCGCGAGCAGGGCTTATTTCGATTT 8880  
DB 69786 ABAACACCTTTCCCGGACCGTCCGGATGCGCGTTCGCGAGCAGGGCTTATTTCGATTT 69845  
QY 8881 AATCGAAAGCTCCACGGATGTAAACCGCAACCGCATCTGGAACCAAAAAACGACACGAT 8940  
DB 69846 AATCGAAAGCTCCACGGATGTAAACCGCAACCGCATCTGGAACCAAAAAACGACACGAT 69905  
QY 8941 GCTAGGGCTATCTTACGAGCCTGTATGACTACAGTCCCGCCCGCCCATCGATTCCCC 9000  
DB 69906 GCTAGGGCTATCTTACGAGCCTGTATGACTACAGTCCCGCCCGCCCATCGATTCCCC 69965  
QY 9001 CTCAGCAACTCGGCAAAATGAGGAGATATTTCCAGAGGTAGTCCGCGCATCTTAGCGGCCA 9060  
DB 69966 CTCAGCAACTCGGCAAAATGAGGAGATATTTCCAGAGGTAGTCCGCGCATCTTAGCGGCCA 70025  
QY 9061 CGGATCGCCTTGA 9073  
DB 70026 CGGATCGCCTTGA 70038

## RESULT 3

AAV19941

ID AAV19941 standard; DNA; 137507 BP.

XX AAV19941;

AC AC

XX 27-AUG-2003 (revised)

DT 03-AUG-1998 (first entry)

XX KSHV long unique coding region and terminal repeat.

DE KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;

KW

KW interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;  
 KW complement-binding protein; glycoprotein; capsid protein IV; infection;  
 KW immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;  
 KW lymphoproliferative disease; leukaemia; splenomegaly; mycosis fungoides;  
 KW HIV immune status; anti-inflammatory agent; therapy; ds.  
 XX  
 OS Human herpesvirus 8.

XX Key Location/Qualifiers  
 FT CDS 1142..2794  
 FT /tag= a  
 FT /product= "complement-binding protein"  
 FT 8699..11236  
 FT /tag= b  
 FT /product= "glycoprotein B"  
 FT complement(117261..117875)  
 FT /tag= c  
 FT /product= "interleukin 6"  
 FT complement(21548..21832)  
 FT /tag= d  
 FT /product= "macrophage inflammatory protein II"  
 FT complement(27137..27424)  
 FT /tag= e  
 FT /product= "interferon regulatory factor 1"  
 FT 28661..29741  
 FT /tag= f  
 FT /product= "protein T1.1"  
 FT complement(58976..60175)  
 FT /tag= g  
 FT /product= "glycoprotein M"  
 FT complement(63412..69915)  
 FT /tag= h  
 FT /product= "glycoprotein L"  
 FT complement(88410..88910)  
 FT /tag= i  
 FT /product= "interferon regulatory factor 2"  
 FT 89600..90541  
 FT /tag= j  
 FT /product= "interferon regulatory factor 3"  
 FT 90173..90643  
 FT /tag= k  
 FT /product= "glycoprotein X"  
 FT complement(93636..94127)  
 FT /tag= l  
 FT /product= "interferon regulatory factor 4"  
 FT complement(111931..112443)  
 FT /tag= m  
 FT /product= "capsid protein IV"  
 FT complement(123808..127296)  
 FT /tag= n  
 FT /product= "immediate early protein"

XX MO9804576-A1.

XX 05-FEB-1998.

XX 22-JUL-1997; 97WO-US013346.

XX 25-JUL-1996; 96US-00686243.

XX 25-JUL-1996; 96US-00686349.

XX 25-JUL-1996; 96US-00686350.

XX 05-SEP-1996; 96US-00587253.

XX 10-OCT-1996; 96US-00708678.

XX 13-NOV-1996; 96US-00728323.

XX 13-NOV-1996; 96US-00747887.

XX 29-NOV-1996; 96US-00758640.

XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Chang Y, Bohensky RA, Russo JJ, Edelman IS, Moore PS;

XX WPI; 1998-130615/12.

XX  
 PT  
 PT  
 XX  
 PS  
 XX

New nucleic acid encoding Kaposi's sarcoma associated herpes virus proteins - useful for, e.g. detecting levels of HHV8 in, and preparation of vaccines for treatment of, HIV patients.

Example 2; Page 135-203; 230pp; English.

This sequence represents the long unique region and terminal repeat of the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known as human herpes virus 8 (HHV8). This sequence contains the DNAs of the invention which encode KSHV polypeptides selected from: (a) viral macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6); (c) viral Irf 1; (d) complement-binding protein; glycoproteins B, M or L; (d) capsid protein IV encoded by ORF65; and (e) immediate early protein encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded by it, and antibodies (Ab) specific for the Kaposi's sarcoma, in body detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body fluids or tissue samples. HHV8 infections can be treated with antitense or triplex forming molecules or agents that bind specifically to the protein. Ab may be used for prophylaxis or treatment of HHV8 infection, while the protein can be used in protective vaccines. Ab may also be used to differentiate between lymphomas, and HHV8 may be implicated in many other lymphoproliferative diseases such as lymphomas, leukaemia, splenomegaly and mycosis fungoides. Cells and animals containing the nucleic acid are useful for drug screening. HHV8-derived peptides can be used as targets for antiviral drugs, e.g. dihydrofolate reductase gene can be inhibited with methotrexate. These can also be used to determine the immune status of a patient infected with HIV. HHV8 derived protein viral MIP III may be used as an anti-inflammatory agent for, e.g. treating rheumatoid arthritis. This sequence is stated as containing 81 open reading frames. (Updated on 27-AUG-2003 to correct OS field.)

Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 U; 0 Other;

Query Match 17.7%; Score 1607.6; DB 2; Length 137507;  
 Best Local Similarity 51.6%; Pred. No. 0;  
 Matches 4481; Conservative 0; Mismatches 3869; Indels 342; Gaps 21;  
 Qy 1 ATGAGAGTTTCGTCGGATGGACCAACACGTCGACCAAAATCCGGGGTTTCATCTTGAAAC 60  
 Db 64892 ATGACAGCTCGGAGGGTGCACTGATATGGACGAGCCAGCCCGGGTTTCATCTCAAC 64951  
 Qy 61 ATGACCTCGGATGCCAAAGTCAGGGGTTCGCGGATCAGTCAGTCGCTGCAAAATATA 120  
 Db 64952 ATGACATCGGATGCCAAAGTCAGGAGCGTAGTGGAAACAGATTGACAGGTTGTCAAAATATC 65011  
 Qy 121 ACTACAGCCACCGGAAATGGGTTGGTACGACCTGGGCTTGGATCCGGCTGAAGACTCC 180  
 Db 65012 ACTACCTGCCCGGAGATGGGCTGGTATGACCTAGAGTTCCGATCCACTGGAAGACGAA 65071  
 Qy 181 GGGCGGTTCTTCCGCTTTACCGTTTCTTAATTACGGGAACGTCGTGGTGGGAAAGT 240  
 Db 65072 GGGCCCTTTCTCGCGTTTCGGCATACGTAATAACGGGACCTGCGAGAGCGGGGNAAGC 65131  
 Qy 241 ACCAGCATATCGGCCCTGTATACCAAAATTTAAATGCTGTATCAAGGCGGACCAACCATTA 300  
 Db 65132 ACCAGGTTATCGGCCCTATCATCAGAAATCTCAACTGCCTTAATTAACGGGGGCTACAGTGGA 65191  
 Qy 301 GCGGCACAGCACTATCGGTCGCTTAAGAGGTTCTGTCCACAGATCTTCAGGCGCTTTT 360  
 Db 65192 GCGGCACAGCACTTTTCAGGGCTTTAAAGTCTCTACTGTCCCATATATATACACGGCTTC 65251  
 Qy 361 GGCTTTAAGCGCGACACATCAATATAGCCGTGAGAAAGCTCATCAGACCGGACCGGTA 420  
 Db 65252 GGATTCAGAGCAGACACATTAATATCTGCCAGAGGAAAGTGCACCAAGGTAACCTCAGTCC 65311  
 Qy 421 TCCATAGACCAATTCAGCAACAGGAGCTATCGAAGTATTGGCGGTTTATAGTGACATTT 480  
 Db 65312 TCCATCGACCACTCCAGAGATACGAGCTGGCTAGGTACTGGCCAACTGTCCACGATATT 65371  
 Qy 481 ATGAAAGAGTTATGCGGAAACCAACCCATGCGCATGTACGGGACTATATCCACCGCAAT 540  
 Db 65372 ATTGAGAAATTTATGCGCAAGAAACAAAGGGGCGAGTATAGTCCCTCTCTCTCAAGCGCT 65431

QY 541 TTGTAAACCTCTCGAGAAATACCGGACCGGTGTTTATGACCTTCCATATATTTGTAATC 600  
Db 55432 TTCAGACTCTTTTCCGTATGCGTGGAGCCAAATTTGTGGAGTAACATATTCGTGATA 65491  
QY 601 GACGAGCGCGAAACCTCTCTCTTCAATCTACCAACCGTGTGTTTCTTTACTGGTTC 660  
Db 65432 GACGAAGCTGGAACCTCTCTCTCCATATTTTGAACGCGCGTGTGTTCTTCTATTGGTTT 65551  
QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGCGGCGGTTCCTGTCATAGTATGC 720  
Db 65552 TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTCCGCTGCTGCATAGTCTGC 65611  
QY 721 GTCCGTTCCCAACGCGAGACAAACGCGTTCAGTCAACTTACCAACCGGAGCGCAAAAG 780  
Db 65612 GTGGGCTCTCCACCCAGACGAGCGCCCTTTCAGTCCGCTCTTCAACCAACGCGAGCAGAGA 65671  
QY 781 ACGAGATATCGTGTGCGAGAACATCTTACATTTTATGATAGGAGAGGTCGTATCT 840  
Db 65672 AACGAGATATCGCTGTGATATGTGTCTACCTTCTTATTTGGAAACGTAAGGTTGCA 65731  
QY 841 GAGTACGTACACTTGGAGAGAAACGTGGCGCTGTATTAAACAAAGCGCTGCACCGAT 900  
Db 65732 GATTATATTAGCTGGAGAGAAATGGGCCCTTATTATAACCAATTAAGCGCTGTACGGAT 65791  
QY 901 CTGCAGTTCCGACACCTTCTAAATAATTTTAGAGTATAATCTTCCCATTTCTGACGAGTC 960  
Db 65792 CCCAGTTGGTCACTTCTGAGACCTTGAATATATCTAGACATATCACAGAGTTA 65851  
QY 961 ATGAGTTACGTAGACAGATTGTCTTCCCTAAAGTAAGATTATGATCTTTTAGAATAC 1020  
Db 65852 ATGAGCTATATAGATAGATTGTGTGTTCCGAAGTAAGATTCTGACCGCTCGAGTAT 65911  
QY 1021 ATTGGCTGACCCGACCTTTTGTGTCATAGCGAGTAAAGCGGTATCTAACAACTTA 1080  
Db 65912 GCAGGTTGGAACAAGCTTCTCACTCAACGAGGTGAAGTCTTTCTGGCAACGCTG 65971  
QY 1081 CACACATGTCTAAACGTTAGGGGGCGATAACGAGGACACGAAGTCTTTACTGTCCCGTG 1140  
Db 65972 CACACCTGCTGTGAGTAATAGGATGCTGTGTCCACAAAGCTTTTTCACCTGCCAGTG 66031  
QY 1141 GTGTGCGAGGTGTTTGTGAAGCGTTTGGAGAAATACAGCGGCGCTCACTTCAACCACTA 1200  
Db 66032 GTCTGTGAGGTGTTTACAGAGCAATTTGAGAGATACAAAGCGCGGTAGCGCTTCAACAC 66091  
QY 1201 CTCACCGTACCGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAAATTTATTCGAGTTT 1260  
Db 66092 ATGACTCCCAAGAAATGGGTAAACAAAAATCTTTTCAGGCTAAGTAACTACTCGAGTTT 66151  
QY 1261 GTGACACGAGCATGTCCATAGTCGCCACGGAATCCACTGAAGTTTCGACCCAGGTTACC 1320  
Db 66152 GCTGATCAGACATGGCTGTGTTGGACCTATATACAGACCGGTCACACAGATCACCC 66211  
QY 1321 TTTATCACCAAGTTGTTAAACAGCCACGTGTCTTTTAAATGGAAAAACAAAAAATGC 1380  
Db 66212 TTCGCCACTAAATTTGTCAAAAAACAGCTATGCTTACCTTACTGGAAAGACCAAAAAATGT 66271  
QY 1381 ATATGTGGTTTCAGGTTAGCTACTTTTGGTTTCAAAAGAAATCTTAGACAGCGAACTCTTC 1440  
Db 66272 ATATGCGGTTTACCGGTATATCAACAAAGATTCAGTCCATCTTAGACGCGGAGCTATT 66331  
QY 1441 GTGAGACGATTCGCAAGCCGTCGGAATACGTATACGTTTCTTAAACACGCTACTG 1500  
Db 66332 ATCGAAAGTCATTGCAACGATAACCCCGCTTATGTGTACAGTTTCTTATGATCCCTGCTA 66391  
QY 1501 TACAACGCCATGTACTCTTTCAGCTACGCGTACGCGTGTAGTTCACACGAGAAATATCTG 1560  
Db 66392 TATATGTCATGTACTCATTTTACGCGCAGCGGTTGAGCAGGGGATGAAGAAATTCCTC 66451  
QY 1561 CAAGACCTAAATTTGACCCCTCCCGGCGCTCTTGGCAACCGGGCGGTGTAGACCTTCAA 1620  
Db 66452 AGGACCTCAGGGAACTCCCGGTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 66511

QY 1621 ACGGTTCTGTGAGAGTTAAACCTTGGAGACGACATCTTCTACCACTGTGTGCTCCCGG 1680  
Db 66512 GTTCTGGGCGAGGAGGGGACACAGATG-----CCTTCTACTCTACCGCAGCTCCCA 66565  
QY 1681 CCACCCCGGGGTATCACTCCCTCCAGGTTTGTGTGACACGCTACTGCGCCCTTAAAGGAC 1740  
Db 66566 CCACTCCCAACCCACGCGGCTCTTCAACACACTGTGTGCTTATTACTCCGCGGCGCAAGGAA 66625  
QY 1741 GTGTTCCGCTCCAGAAATAAGGTGGGTGTGCTGCTTTTGGCGGGAGGTTTGAAGAGGAA 1800  
Db 66626 CTATTTCTGCAACAGCTGGCCCTGTGCACGCCGACACTTTTGTGTACAGTTCTTCTACTCC 66685  
QY 1801 ACGTTTCCGCGTTTACGTTTAAACATGTGTAAAGGAGCGGAGTTGACTTTGTCTCCCT 1860  
Db 66886 GATTTTTCACACGTTTACGTTGAAACATGTGTGCGAGATGGCGTGACATTGTGTCCAAT 66745  
QY 1861 TCAGAACGTTCTCAACGCGCTGTTGGGCTTGTGCATCGACCGTTGCAATCGCTATAAATAAG 1920  
Db 66746 TCCCCCGGCTCCACGCTGTAGTGGCATACGCTATCACTATAGACACCTATATATATCCAG 66805  
QY 1921 GGGTACACGTTTTCACGTTAGCGTGTGCTGTGTCAGGGTTTGCACCTCACTGATGATAC 1980  
Db 66806 GGATATACGTTCTCCAGTGAGATTTCGCGCTGCCAGAGGACAGCGCTCCAGCAGGAC 66865  
QY 1981 CTCAGGAAGAGATGCTCCCTGCTGTCAGAGCTCTTAGCGGTTTATCGGTCGCTTA 2040  
Db 66866 CTGCGCAGAAAGATGCTCCCTCATAGTTGTCCAGACTCATCGGGTTCTATTGCTGCTG 66925  
QY 2041 GAGATTAACATAACCAATTTGACCGAAACCATGAGGACCGGAGCGTTTTCOAAGTGTGC 2100  
Db 66926 GAAATAACGCTCAACCAAGATGACAGAGACCTCGAAGGTGGGACCGTGTTTTAACTATGT 66985  
QY 2101 TGTGCGGGGACTATGCGGTTCAGTCAAAATTTAGCCATGACCATCGTAAGGCACAGGGA 2160  
Db 66986 TGTGCGAGGACTATCGGTATCAGTTCTTAATCTGCTATGACCATAGTAGAGGACAGGGG 67045  
QY 2161 ATGTGTTGAGGAGTAGCGGTAGTATTGTTGGTCCCAAGAAAGCTCCAGACAGCCAC 2220  
Db 67046 GTTTCATAAGTAGGGTGGCCATATGTTTCGGCAACCAACCGCAATATCAGAGCCAGTCTA 67105  
QY 2221 GTGTATCTAGCAATATCAAGGCTGTTAACTCAAAATTTTGTGCTATGACAGACACCC 2280  
Db 67106 GTGTATGTGGGTATTCAGGGCCATCGAGCTCGTTACTGTGTATGTGACAGTAACTCC 67165  
QY 2281 CTTAAACCCCTCTCAGAGAAACAGTTCGATAACCTCCGCGCAAGATATATGTCGCGCC 2340  
Db 67166 CTTAAGCTAATGACCGCGGTGACGCGCTAGTCCCATCTCTCAAGTACATCATCAAGCC 67225  
QY 2341 CTCACAAACCCAAACCAACCTCTCATCTACTAA----- 2373  
Db 67226 CTATGCAACCCCAAGACTCTCTGATCTACTGACCGGTACCCCTCTCTTAGGACACTGAT 67285  
QY 2374 -----ATAATAAAAAACAAGAAATGGAATAACATTGTG 2406  
Db 67286 GTGTTTGGGAATAAAGCATGAGACTTGCACCTATAATGGTCTGTATTGACACCATTTCT 67345  
QY 2407 TTTTATTAGTCCCAACCAACCGCCAGGATAGTGTCAATTTCCACACACCGGGGGGTGG 2466  
Db 67346 TTAATTTAGTCCACCGCCAGGCTTATATGACCGTTTCCACAGAGGTTGCGTGG 67405  
QY 2467 CCGCCATAGTTTGAACACAGAAAGCCCGGCGCTCTGTTGGTTAGAAGTAGGCC----- 2521  
Db 67406 AGGCCAGGATGCGGGTTGGGTGCTGACCTGGACCCCGCGGTAGTTGTGTTCTCTGATG 67465  
QY 2522 -----GCCCGCGGTGTCGAGGTGATACGCTCTCTTAGGGCGCTGGGTT 2564  
Db 67466 AATCAGGTGGCGGGAAGTACTGGGAGATTGGTTGGGAGGTGACCTTTGTGCTCGAG 67525  
QY 2565 GTACCAACATAAATCAGTCTGTCATCACCGTTGCAAGTCCCTGCGTACGCGGCGCTTTG 2624  
Db 67526 GAGACACGATCAGCTTCACGCGGAGGAGGCTCTCTGCTGTGTCTCTCCCGGAGGATA 67585  
QY 2625 CAGGCGGACGACACCGCGGCTTTGTAAGTGTATCTG-----CAT 2664



|    |       |   |       |
|----|-------|---|-------|
| Db | 67586 | TAATTATCATCGGACGGCCACTGCTTTGGCGCTTAAGTTTGGTTGCTCTGGGACGCGCACCA      | 67645 |
| Qy | 2665  | CGTCGTGCTGTCTGAAGATGAACACAGGTCAATCACATCATAGTGGCGCCGTGCCTCAT         | 2724  |
| Db | 67646 | CATCCTCGCTACACAGAGGAGCGGTAGACTGCCTTTTGGCTTCTTGCGCCACGCTCCATGA       | 67705 |
| Qy | 2725  | CGCGCTCTGATTCTGTCCTCCCATGCTCAACAGAGCCCTCGGTCTGTCACACTTCTGTCGTCATCTG | 2784  |
| Db | 67706 | GCCCGATTCTCTGACTCATATCTTCCCTTGGTCTTCTCCGTCTCTCTCGACAGAGGTG          | 67765 |
| Qy | 2785  | AAATAATAATAGTCTCTTGTCTCGAGTTGATCCAGCAACAAACAGTTTCTTTCATCGG          | 2844  |
| Db | 67766 | GCTGGTGGAAAAATAGCCCGCGTCCGTAAACCGCGCCTCATTTGTCACGTCGAGAGT           | 67825 |
| Qy | 2845  | AAGACCATCCGCCCTCTGCGCGTGTATCTAGATATATCACTGACAGAGCGCGGACAGCCC        | 2904  |
| Db | 67826 | TGGAATGTCTATCGCTATCAGAGTCCGATGTCAAGTCCACCATCCGGTGGGTGCGCGC          | 67885 |
| Qy | 2905  | G-----  | 2905  |
| Db | 67886 | GCAGGGGGGCCACGAGGGCCCTTCATCAGGCTCGCTGTATGGTGAATTTGTGTTCCAG          | 67945 |
| Qy | 2906  | -----TGCGGAAAGCCGTACGTTCTCCGCCACTTTCATCATCGTGT                      | 2946  |
| Db | 67946 | GTACACTATTTCTGGAAGCAGGTGAATCCGATGCCCGGTCCAGTGTATCCGCCCA             | 68005 |
| Qy | 2947  | CCTGTAGTCACTATAAAAAAATCGTCTTCTGCTCTGTATCAGATGATCGGGGTGCGC           | 3006  |
| Db | 68006 | TGGTTTCCAGGATAGCAACCCCTCGTCTGAAGGTGAGAGCCGACGAGGGGAAATC             | 68065 |
| Qy | 3007  | GGTGTGGCGAGTAACACCGCAACGGGCGCG-----GACCACCAAAACGCG                  | 3053  |
| Db | 68066 | CGTCATCTGACTAATCCCATCCATGGAGCCTCGGACTCCGCGGTCTGTTGAATGC             | 68125 |
| Qy | 3054  | GTACAGTGGCGGGCGCGCTCGGTGGTGGTGGCACTGTGCGAAGTAAGACGGTAGGT            | 3113  |
| Db | 68126 | GCACGGCGCCCTACCACTGCTACCGTTTGGCGGTATGGGCGCGCTGCGCCAGAGGCC           | 68185 |
| Qy | 3114  | AGCGCGCGCTCTGGTTTAAAG-----ACTCACCGTCCCGTG                           | 3150  |
| Db | 68186 | TGGGCGCAAGTGAGATAAAGTTGAAAAAGTCTGACGGGTACCCCTCTGGCTCGTCT            | 68245 |
| Qy | 3151  | CGTGGGATCTATTTAGTGGGGAGGAGACCGGCGGTTCTGTTCTTTCATCCCATCAC            | 3210  |
| Db | 68246 | CCTCTGAAACATCGTCAATTTCTTCTTCATCTTCATCTCTCATCTCTCGTCATATTCAG         | 68305 |
| Qy | 3211  | TGTAGTCGCTGTCTCCGAATGAGAGTCGGTTTCGTCTGGCGTTTGGAGTAGTGGAG            | 3270  |
| Db | 68306 | ATTTCGCGCTCGACTGATCCGGGGATATCTGTAGATCCAGAGGGTTGCTGGCGCGATG          | 68365 |
| Qy | 3271  | GGGAATCTTGGTCAAAAACATCTCAGGTGCGTCAATCAATCACTGGGGGGCTCCGT            | 3330  |
| Db | 68366 | GGGTGTCCTCGCGAAGACGTGCTCTGGGCGAGACATATCTATCACCGTGGGTCCAGCAT         | 68425 |
| Qy | 3331  | ACCGGGCGGTATCCGGCAACCCCGTACGGGA-----CCGCGGCAACAAAAAC                | 3378  |
| Db | 68426 | AGCGCGCGCTGCCCAATCTCGAATGATGAAGAGGTGGAGTGGGAATATGAAT                | 68485 |
| Qy | 3379  | GTCCTTGTCTTTTTTCTAGTGGCCCGGGAAACGCGAGCATCTGGGAGTGGCGCGC             | 3438  |
| Db | 68486 | TACGGGGGGTCTCTGCGAGCGCTCTTCAATTGGAAGCATCTCTCTTTCATCGTGTG            | 68545 |
| Qy | 3439  | GGGTCTTTCAGCGGTCCGATAAGAAC-----ATAGCCATGGCC                         | 3476  |
| Db | 68546 | TGCTAGACAGGTCTTCAAAACATCGCCATGCGCTTGTACGGGTTGACCGCTAGGGC            | 68605 |
| Qy | 3477  | GAACGTTCACTGTAAAAAGCAGCTGCGCTATCCCGGATCAGTTTAATTTCCAGTCAACGG        | 3536  |
| Db | 68606 | GGAAATTACAAAGCACAGATTATGCTTTTACTGCTCCAAACAGGCCCCAGTCCACAG           | 68665 |
| Qy | 3537  | CACCCCTCGGATTTCTGAACATAGTATCATTCGCAAGTTTAAATGGTTGAGCGCTAGAA         | 3596  |

68666 TCTCAGCCGGTGGCGAGTCAATAGTCGTTGGCTAGGTTAAAGTAGATTACAGCCCTGGA 68725

3597 ACTGGCCGACGAGAAACCGGTTGCCGCTGGCGGTAGTCAGGGGAGACGATGCCGC 3656

68726 ACCGAGGCCATCGCGAGTGTGGGCCACCAAGAGAG--GCCAGCGAGATGATGCTGG 68782

3657 CATTGAGAGCAGATGCTCGTCCGTTATCAGCGTCCGCTTGCTATAGCTCTTCCTC 3716

68783 CCGTAAGCACCAAGTGTTTCTGTGCGTTTATGAGCGGAGTTCGTGCAATGGCTTGGCC 68842

3717 CCCACAGCATAAAAACGCACTCTGAAGCTTGGCGAAGAAAGGACGCTTATATGTAGTGG 3776

68843 CCCACAGAGAAAAACGCAATGTTCTAACTTTGAGGATATGCTACTGATGATGAACAAC 68902

3777 TGAACCAATCCAGCCAAAGTGGCGTGTGACCCCGCTTCCCTCTTTCCACCGTCAGAA 3836

68903 TGAACCAATCCAGCCAAAGTCCCTCGTGTGAGCCGGCCCTCCCTCTCCACCGTCAAA 68962

3837 TGGTGTTTAAAGCAGAACTCCCGTTTTTGCCCAAAGTCTAAGCAACCGTGAGAGAGAG 3896

68963 CTGTGTTTAGTAGCAACACACCTGGCGAGCCAGCTGTGAGGACACCGCTGGGAAGAG 69022

3897 CGGTGAACCCAGGTACAGTATGGCTATCTCTTTAAAAATATTTTGAAGCTTTGAGGTA 3956

69023 TACTGAATTCGGGACGGAAGCCTTAGCTCTCTAAGATGTTCTCAAACTGGGTGAA 69082

3957 TACTGTAGTCGGAGCAACGCTAAAGCTAGTCCGGTGGCTTGAACGCGATGCTAAGAT 4016

69083 COTGACATTTGGGATCCACACTAAAGCCAGGCCAGTACTGTGGCCCTTGTGGTACGG 69142

4017 COTGGCCAAAGGATCAGCACTTTAATATCTCTCCGGTTTCGCAACAAATACGACCAACCATAA 4076

69143 COTGGCCCTAAGATCACCACTTTAATATCTCTCGATCGCAGAGTGGGACCAACACATCA 69202

4077 TTCTATCTATCGTGGATTAATCAGCGTTAGTGTCTCATATCCATACAGCGCTTCAGCA 4136

69203 GCTTGTCTGTGGGGATACACTGTGGTGGTTAGCCCTAAGTTCCCGAATCTGCTGAGCA 69262

4137 GAGCCGCGAGCTTCGTGTTTTAAAAATGGAGAAAGATTTAAAAATTCAGACCATGATCGC 4196

69263 GCGAGAGCAGTTTCTGTTTTCAGAAATGATGAGAGGCTCAGAAAGGAAATCCACTTAGTG 69322

4197 TAAGTAGCA-----AGTTTGTGCGAGGGTCTCTCAACACT 4234

69323 CCAGTAACAGATCCCGTGTGCCACCCCTGACTGATGGATAGGTTGGCCCTTAAGACCG 69382

4235 TCTGTGTGACATTTACTTCCAAACTATTGTTTTAAGCCAAACCCCTCCATAAATCCCGCGTC 4294

69383 TCTGTTGCAACCATGCGTCCATGTTGAACTTATTTTCCCTTTTGACCTGCGTGGCTCTC 69442

4295 TGGTAATTACTTTACTGTATAGTGAACATTTTGACTCTGTCTGAGAAATCCACAGAG 4354

69443 CGGTGCTGCTTTTAGCCCGAGTCTGACTTCCGCTTAACAGAACCTGTGCGGTTTCAAGGCC 69502

4355 GTGGCCGCGTATAAGCAGTTCCGCTTTCACTACCGTTTATGTGAAAGT-----CT 4404

69503 TTTCCACGCTTATATAATATGTTTACGTTGTGAATAGAGCTATCTGCAGTGGTGGC 69562

4405 TCAAACTCGGTAAAACTCATCAGACTATATTGCAATATAATTCAACAATTTGAGCTCT 4464

69563 TTAACAACCTACAGTATAGGCGGTCAACCTCGTTGTAAATATCCACAAACCTCAGTTT 69622

4465 TCTCTGGGACACATGTAATTTAACTTTGTACAAAATCTCTTATAAAAACGCCCTTAAGTGA 4524

69623 TCTTCGAGCGCCACGAGGCCCAATCTTCGAACGACCGGACTATAAAATGAACCTCAGATTA 69682

4525 AATCCATTTACATATTTCTGTGGTAGAGCGATCAGCTTTATATCGCAATTTGCTCTGA 4584

69683 AACCCATTCAGCATGTTTCCACGGTAATGTGCGCTGTTTTGCTTCGAGAGCTTGGCTATA 69742

4585 CACAGCCCTTGACATAAAGCTAAATTTGTTAAATGAAAGAGCAATCTCAAAACGGCTCC 4644

69743 CAGACCCCGTTGAGTCAATTCGATCGGCGAAGTGGATGATGAGTGAGCCGCAAGAAACAC 69802





|    |       |   |       |
|----|-------|---|-------|
| Db | 71961 | GTCTTACCAAGAAAACTTTTACCACTGCGCCATCCACTGAAAGGAGGACAC         | 72020 |
| Qy | 6824  | AGGTCGTGAAAGTGTCTATCATCTCTTCACTCCCTTTGAGGTTTGGTTACACCG      | 6883  |
| Db | 72021 | GTCCGTTGTGGTGTGTAGGATATCCCTAACTTCGAGCGGAGACGGCGACGTCCC      | 72080 |
| Qy | 6884  | ACATGTGAGATAGGACCAAGCTCAGGGGGAGGTCGAGTCGTAGGGGGTAAACATTCG   | 6943  |
| Db | 72081 | ACAAAATGGGAGAGGCCACCTCTGTGTCAGTCCCGGTCTGGGGTTCTGATTCACGGGC  | 72140 |
| Qy | 6944  | TGGCGTGGGGGTACTCCAGTTTATAGATCTCCGGCAGGACGGTACGAGTTCTTCTGTC  | 7003  |
| Db | 72141 | GCCGTGTGGGGTATTTGAGAGTCAAACTCTCGGCAGTCCCTTAATGAGCTCTCTCA    | 72200 |
| Qy | 7004  | AATCTTACGAGCCCGCATCCACAGGGGAGTACTGAGTTAATAAGTCCCGCAATCTA    | 7063  |
| Db | 72201 | AAACCTATGAGCCAGCGTCCACTAGTGGCAGCATGCCGTAAATAACACCCCTTATCTTG | 72260 |
| Qy | 7064  | TCGTTCTGTAATTTGTGGAGCTCTTTAGACAGTAAACATGTTGTCGACACATCGCGT   | 7123  |
| Db | 72261 | TCGTTGCCAAGTTTGTAACTGCTGACAGGAAATAGCCAAATTCGCCCTAGCCGGGA    | 72320 |
| Qy | 7124  | ACGATGAACCGCTCTCTGTTGTCATGAGCACCTGGAGTATCTTTAGACAGTAT       | 7183  |
| Db | 72321 | ACCAGGTACGGCTCGCTTGTGCGTGTGGACCAATATCTGAATGCTTTTCAAGTAT     | 72380 |
| Qy | 7184  | AGGCCCGTTCCACTCTAAGGTTGCGCTTGTGAGAGCGGATTCGGGCATTTGACGTC    | 7243  |
| Db | 72381 | AGGCTTCTTCAAGGTTTACAGCGGTACGTCGACGTCGGATTTGAGGTTGGCGACGAC   | 72440 |
| Qy | 7244  | AGTGTGTTCAATTCACCGACTATCATGCCAGTGGGGATAGTGGAACTTTGACATATGA  | 7303  |
| Db | 72441 | AGGATATCTAATCTCTGAATCTGATCCAGGACGGGTAAATGATACCTAAACATGG     | 72500 |
| Qy | 7304  | TTTAAAGATGGTCTTCCCGTAATGGCGACTCATGTCGTGTTTTA-----AC         | 7352  |
| Db | 72501 | TTGAACAGGTGATCTTTAAGGGGCTTCTCGATGTCATTTGATAAAATATGACACGCCAC | 72560 |
| Qy | 7353  | GTTCACAGGTGTTAAATAACTCAGAAGTCTTGGCTGACAACTTTTGGGACTCTCG     | 7412  |
| Db | 72561 | TCTCTCTTAGGTAAGAACCTTCGGCGTCTCTGTGTGGAAGCTTCGTCGGCTCTCG     | 72620 |
| Qy | 7413  | AACCGCTCAGGATAGCTTACCAACCGGTGCTTTTATAAAGCAAGCATACGTCAAAA    | 7472  |
| Db | 72621 | ACGAATGAAGGCCAACTTACCAAGTGTGTCTCTTATAAATGACGCATACGAACAA     | 72680 |
| Qy | 7473  | AACATGACCCAGTACCTAAATAGCTCAGAGAGTCTCTCGATGTATGTATCG         | 7532  |
| Db | 72681 | TCTACGATCCAGTGACCTAAATAGAGTGTGGAAGATGTGATTCGATTCGATATGAAG   | 72740 |
| Qy | 7533  | AATGTGCTCCCTGGGTCCCATCTCCGGGTCTATCGCGATCTGAATCTCTTTAACCTGT  | 7592  |
| Db | 72741 | AATGTTCAAGCTTGGTGGCTATGTGTCTGTTTACAGACATTAACATGTTTAACTTT    | 72800 |
| Qy | 7593  | TCTGCTCTACCGGGTTCCCGGTTAAACCCCGGGGCGGCACCTGTAAGTCCCGT       | 7652  |
| Db | 72801 | TCTGCTTTTTCGTGCTCTCGAATGAGGACCAAGCGCGCGGTACACGTGCCAT        | 72860 |
| Qy | 7653  | GCGCAGATGCGCCAGGATCTGTGAGATTTCTACCGAGAGGCGCTTGTGTACCG       | 7712  |
| Db | 72861 | GCGCAGGATCTCCAAAGCATATTTCGATCTCTCAGAGAGATCTTATTCGACAG       | 72920 |
| Qy | 7713  | AAAAGATGTTATAGCTCTCCCTGACGCGGGTGTATACCGCCCGAGTGGCCAGAG      | 7772  |
| Db | 72921 | AAAAGGATTTCTGACAGCGCATGACGCGGGTGAAGCTTCCAGCCATATGTAAGC      | 72980 |
| Qy | 7773  | TTCTGACGAGCTTACCGGAAATGAAGCCAGTGTGGGGGGTGGCGTCTCA           | 7832  |
| Db | 72981 | TACTACGAAATATACACTGAATGAAGGCCAAATGCTTGGGGCTGAGCGGACTCG      | 73040 |
| Qy | 7833  | TATGTTGAGCGGCGCTATATGCTATTCGCACTCGGTCTCTGTCATTAATACCC       | 7892  |

RESULT 4

AAV73803

ID AAV73803 standard; DNA; 35100 BP.

XX AC AAV73803;

DT 17-OCT-2003 (revised)

DT 25-FEB-1999 (first entry)

XX KSHV LUR DNA (nucleotides 35,101-70,200).

KW Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;  
 KW dihydrofolate reductase; LUR; long unique region; vaccine; prophylaxis;  
 KW diagnosis; treatment; HHV8; thymidine kinase; glycoprotein H;  
 KW capsid protein; packaging protein; helicase primase; glycoprotein M;  
 KW viral protein kinase; alkaline exonuclease; virion assembly protein;  
 KW uracil DNA glycosylase; UDG; glycoprotein L; ds.

XX Human herpesvirus 8.

OS US5849564-A.

XX PD 15-DEC-1998.

XX PF 29-NOV-1996; 96US-00770379.

XX PR 29-NOV-1996; 96US-00770379.

XX XX (UYCO ) UNIV COLUMBIA NEW YORK.

XX Bohenzky RA, Moore PS, Russo JJ, Chang Y, Edelman IS;

XX WPI; 1999-069741/06.

XX Kaposi's sarcoma-associated herpes virus nucleic acid - encodes  
 PT dihydrofolate reductase and is useful for treatment, prophylaxis or  
 PT diagnosis of Kaposi's sarcoma.



|    |       |  |       |
|----|-------|--|-------|
| Db | 31526 | CTATTCTGCAACAGGCTGGCCCTGGCACCCGACACTTTTGGTGAAGAGTTCTCTCCACTCC  | 31585 |
| Qy | 1801  | ACGTTTTCGGGTTTACGGTTAAACATGGTCGTAAGGACGGAGTTGACTTTGTCTCCCT     | 1860  |
| Db | 31586 | GAATTTTCAACGTTTACGGTGACACATCGTGGTCGAGATGGCGTGAACCTTTGTGTCCACT  | 31645 |
| Qy | 1861  | TCAGAACGTCCTCAACGGGCTGTTGGGTTTGCATCGACCGTTGAATCGTATAAAATTAAAG  | 1920  |
| Db | 31646 | TCGCCGGGCTCCACGGTCTAGTGGCATACGATCACTATAGACACCTATATATCCAG       | 31705 |
| Qy | 1921  | GGGTACACGTTTTCACGGTAGCGTTTCGGTTCGCTGCAGGGTTTGCCTACTCAGTGAAGAC  | 1980  |
| Db | 31706 | GGATATACGTTCTCCAGTCAGATTCGGCGCTCCAGGAGACACGCGCTCAGCGAGGAC      | 31765 |
| Qy | 1981  | CTCAGGAAGAAGATGCCCTCCCTGGTCGGTCAGGACTCTAGCGGTTTATCGCGTGCTA     | 2040  |
| Db | 31766 | CTCGCAGAAAGATGCCCTCCATAGTGTCCAGGACTCATCGGGTTTCAITTCCTGCGCTG    | 31825 |
| Qy | 2041  | GAGATTAACATTAACMAATTGACCGAAACCATGGAGGACGGAGCGTTTTCAGAGTGTGC    | 2100  |
| Db | 31826 | GAAATTAACGTCACCAAGATGACAGAGCCCTCGAAGGTGCGAGCTGTTTAAACATATGT    | 31885 |
| Qy | 2101  | TGTGCGGGGACTATUGGGTCAAGTCAAAATTTAGCCATGACCATCGTAAGGGCACAGGA    | 2160  |
| Db | 31886 | TGTGCAGGGACTACGATATCAGTTCTAATCTGGCTATGACCATAGTGAAGGCACAGGG     | 31945 |
| Qy | 2161  | ATGTGCTTGGAGGAGTAGCCGTAGTATTTGGGTCCCAAGAACGTCGACAGCAAGCCAC     | 2220  |
| Db | 31946 | GTTTCTACTAAGTAGGGTGCCCATATCGTTGGGCAACCAACCGCAATATCAGAGCCAGTCTA | 32005 |
| Qy | 2221  | GTGTATGTAGCAATATCAAGGCTGTAACTCAAAATTTTGGTCATGGACAGCAACCCC      | 2280  |
| Db | 32006 | GTGTATGTGGTGTATCCAGGGCCCATCGACGCTCGTTACCTGGTAAATGGACAGTAATCC   | 32065 |
| Qy | 2281  | CTTAAACCCCTCCTCAGAGAACAGTCGATTAACACCTCCGCGACAGCATATAGTCGGGCGC  | 2340  |
| Db | 32066 | CTTAAGCTAATGACCGCGGTGACGCCAGTCCCATCTCAAAAGTATCATCAAAAGCC       | 32125 |
| Qy | 2341  | CTCCACAAACCCAAACACAAACCCTCATCTACTAA                            | 2373  |
| Db | 32126 | CTATGCAACCCCAAGACTACTCTGATCTACTGACCCGTACCCCTCTCTTAGGACACTGAT   | 32185 |
| Qy | 2374  | -----ATAAATAAAACAGCAAAATGGAAAAATTTGTG                          | 2406  |
| Db | 32186 | GTGTTTGGGAATAAAGCATGAGACTTGACACTATATATGGTCTGTATTGACACCAATCTTT  | 32245 |
| Qy | 2407  | TTTTTATTCAGTCCAAACACGCGCCACGGATAGTTGTCAITTTCCACACACCGGGGGGTGG  | 2466  |
| Db | 32246 | TTATTTATCAGTCCAGCCACGCCAGTTATATGCACCGTTTCCACACAGGGGTGGCGTGG    | 32305 |
| Qy | 2467  | CGCCCATAGTTTGAACACCAAAACGCCGGGCCCTCTGGGTTAGAGTAGGCCCC          | 2521  |
| Db | 32306 | AGGCCAGGATGCGGGTTGGGTGCGTGCACCTTGGACCCCGCGTATGTTGTCTTCTGATG    | 32365 |
| Qy | 2522  | -----GCCGCCCGTGTGAGGTCATACGCTCTCTTAGGCGCCCTGGGTT               | 2564  |
| Db | 32366 | AAATCGAGTGGCGGAAGTACTGGGAGATTTGGGAGGTGACCCCTTTGTCTCGACG        | 32425 |
| Qy | 2565  | GTACCAACATAATCACGTCGTATCATCGGTTGAGTCCCTCGGTGACGCGCGCTTTG       | 2624  |
| Db | 32426 | GAGACACGATCATCGTCAACGCGGACGAGGGCTCTGTCTGTCTCATCTCCCGAGGATA     | 32485 |
| Qy | 2625  | CAGGCGGCACGAACACCGCGCTTGTAACTGATCT-----GCAT                    | 2664  |
| Db | 32486 | TAATATTACGGACGCCACTGCTTTTGGCGCTTAAGTTTGGTTGTCTCTGGCAGCGCACCA   | 32545 |
| Qy | 2665  | CGTCGTGCTGTCTGAGATGAACACAGGTCAATCATATCATAGTGGCGCCCGTGCCTCAT    | 2724  |
| Db | 32546 | CATCTTCGTATACAGAGGAGGCGGTAGACTGCTTTTTCGCTTCTTGCGCCACGTCATGA    | 32605 |
| Qy | 2725  | CGCGGCTCTGATTCTGTCCTTCTCAACAGAGCCCTCGGTCTGCTCCACTTCGTCTCATCTG  | 2784  |

|    |       |   |       |
|----|-------|---|-------|
| DB | 32606 | GCCCGATTCTCGACTCAATATCTTCCCTCTGGTCTCTTCCTCGCTCTCTCGACGAGGGTG    | 32665 |
| QY | 2785  | AAATAATAATAGTGTCCTTGTCCTGCAAGTTGATCCAGCAACAACAGGTTCTTTCTTCATCGG | 2844  |
| DB | 32666 | GCTGTCGGGAAAAATGCGCGCGCTCGGTAAACGCGCCTCATTTGTTCACTGTCGGAGAGT    | 32725 |
| QY | 2845  | AAGACCATCCGCCCTCTGCGGGTGATCTAGATATATACCTGACACAGCGGGCGAGGCC      | 2904  |
| DB | 32726 | TGGAACTGTTCATCGCTATCAGAGTCGATGTCAGGTGACAGTCCGGTGGTGGCGGCC       | 32785 |
| QY | 2905  | G-----  | 2905  |
| DB | 32786 | GCAGGGGCGCCACGAGGCGCCTTTCATCAGGGTCGCTGTATGGTGAACCTTTGTGTTCAG    | 32845 |
| QY | 2906  | -----TGGCGAAAGCCGCTCAGCTTCTCCGCCACTTTTCATCAGTGT                 | 2946  |
| DB | 32846 | GTACACTATTTCTGGAAGCAGGTGAAGTCCGATGCCCGGTCCAGTGTATGCCGCA         | 32905 |
| QY | 2947  | CTCTGTAGTCATCTATAAAAAATCGTCTTCGCTCTCTGATTCAGATGACTGGGGTGGCG     | 3006  |
| DB | 32906 | TCGGTTCACAGATAGCAACCCCTCGTCGTCTGAAGGTGAGAGCCACACAGGGGAAATC      | 32965 |
| QY | 3007  | GGTGTTGGCCAGTAACCAACCGCAACGGGC-----CGGGACCAACCAACGCG            | 3053  |
| DB | 32966 | CGTCATCTGACTAAACCATCCATGACGCTCGGACTCCGCGCTGTCGGTGAACCTGC        | 33025 |
| QY | 3054  | GTACGTGCGCGCGCGCGCTCGGTGTGTGTGGCACCTGTGCAAGTAATGACCGTGTGT       | 3113  |
| DB | 33026 | GCACGCGGCGCGCTACCACTGTCTACCGTTTGGGGGTATGGCGCGCTCTGCCAGAGGCC     | 33085 |
| QY | 3114  | AGCGCGCGCTCTGTGTTAAAG-----ACTCACCGTCCCGTG                       | 3150  |
| DB | 33086 | TCGGGCGCAAGTCAGATAAAGTTGAAAAAGTCTGCAGGGTACCCCTCTGGCTCGTCTT      | 33145 |
| QY | 3151  | CGTGGCATCTATCTAGTGTGGGAGAGACCGCGGTTTGTCTTCTTCATCCACATCAC        | 3210  |
| DB | 33146 | CCCTCCGAACATCGTCATTTTCTTCTCATCTCTCTCTCTCATCTCTCGTCATATTAC       | 33205 |
| QY | 3211  | TGTAGTCGTGTTCTCCGAATGAGAGTCGGTTTCGTCTGGCGTTTTCGAGTAGTGGAG       | 3270  |
| DB | 33206 | ATTCCGCGCTCGACTGATCCGGGGATATCTGTAGATCCAGAGGGTTCCTGGCGCGCATG     | 33265 |
| QY | 3271  | GGGAATCTTGGTCAAAAAATCCTCAGTGTGAGTCATGTCATCTACTGGGGGCTCGT        | 3330  |
| DB | 33266 | GCCTGTCTCTCGGCGAGAGCTGCTCTGGGCGACATATCTATCACCGTGGTCCAGCAT       | 33325 |
| QY | 3331  | ACCGGGCGGTACCGCAACCCCGTACGGGA-----CCGCGCAACCAAAAAAC             | 3378  |
| DB | 33326 | AGCGCGCGCTCGCCAAATCTCGAAGTGATGAAGAGGTGGAGTGGGAATATGAAT          | 33385 |
| QY | 3379  | GTCTTGTCTCTTTTCTAGTGTGCGCGGAACTCGCAGCATCTCTGGAGTGCCTCGCG        | 3438  |
| DB | 33386 | TCACGGGGGTGCTCTGCGAGCGCTCTCTCAATTTGAAGCATTTCTCTCTTCATCGTGTG     | 34445 |
| QY | 3439  | GGGTCTCTTACGGGTTCGATAAGAAC-----ATAGCCATGGCC                     | 3476  |
| DB | 33446 | TGCTAGACGAGGTCTCTCAAAACATCGCCATGGCCCTTGTACGGGTTGACCGCTAGGGC     | 33505 |
| QY | 3477  | GAAACGTTCACTGTAAGACCACTGCGCTATCCCGATCAGTTAATATCCAGTCAACGG       | 3536  |
| DB | 33506 | GGAATTTCAAGCACACGAGTTATGCTTTTCTGCTCCAACAGGCCCCCGTCCACAG         | 33565 |
| QY | 3537  | CACCCCTCCGATCTTGAATAGATGATCATTCGCAAGTTTAAATGGTTGAGCCTAGAA       | 3596  |
| DB | 33566 | TCACGCGCGTGGCGAGTCAAAATAGTCGTTGGCTAGGTTAAAGTGATTAACGCCCTGGA     | 33625 |
| QY | 3597  | ACTGCGGCGCAGGAGAACCGGTTGCGGTGGGGGTAGCCAGGCGGACACGATGGCGCG       | 3656  |
| DB | 33626 | ACCGAGGCCATCCGAGTGTCCGCCACCAAGAG-----GCCACGGAGATGATCTCTGGG      | 33682 |
| QY | 3657  | CTTTGAGGACGAGATCTCTCTCGCTTATCAGCGTCGCGCTTGTCTATAGCCTTTCTTC      | 3716  |
| DB | 33683 | CCGTAAAGACCAAGTGTCTCTGTGGTTTATGAGCGGAGTGTCTGTCAATGCGCTTGGGCC    | 33742 |

QY 3717 CCCAGCAGTAAACACGACCTCTGAGCTTGGCAGAAAGCAGCTATTATATAGTCGG 3776  
 Db 33743 CCCACAGGAGAAAACGCAATGTTCTTACTTTGAGGATATGCTACTGATGAACACTCG 33802  
 QY 3777 TGAACCAATCCAGCAGCAAGTTGGCGTGTGACCCCGCTTCCCTCTTTCCACCGTCAGAA 3836  
 Db 33803 TGAACCAATCCAGCAGCAAGTTGGCGTGTGACCCCGCTTCCCTCTTTCCACCGTCAGAA 33862  
 QY 3837 TGGTGTTTAAAGCGAAGTCCCGTGTGACCCCGCTTCCCTCTTTCCACCGTCAGAA 3896  
 Db 33863 CTGTGTGTTAGTAGCAACACACCTTGGCGAGCCAGCTGTGAGCCACCCGCTGGAAGGAG 33922  
 QY 3897 CGGTGAACCCAGGTACAGTATGCTATCTCTTTAAATATTTTGAAGCTTGGAGTA 3956  
 Db 33923 TACTGAATTTGGGAGCAAGCTCTAGCTCTCTTAAAGATGCTTCTAAACTGGTGGAA 33982  
 QY 3957 TACTGTAGTCCGAGCAACGCTAAGGCTAGTCCGCTGCTTACCGCGATGTTAGGAT 4016  
 Db 33983 CCTGACATTTGGGATCCACATTAACCGCAGGCGAGTAGTTCGCTTGTGTCGGGT 34042  
 QY 4017 CCTGGCAAGGATCAGCACTTTAATATCTCCGCTTGGCAATAGACACCAATCA 4076  
 Db 34043 CCTGGCTTAAGATCACCCTTTAATATCTCTGATCGCAGAGTGGGACCAACATCA 34102  
 QY 4077 TCTATCTATCGTGGATAAATCAACGCTTACGTTGCTCATATCCATAACGCGCTTCAGCA 4136  
 Db 34103 GCTTGTCTCTGCGGAGTACACTGTGCTGCTTACGCTTAAGTTCGCAATCTGCTGAGCA 34162  
 QY 4137 GAGCGCGAGCTCTGTTTAAATAAGGAGAGATTTAAATTTCAAGCCA 4188  
 Db 34163 GCGAGAGCAGTTCTGTTTCAGAAATGATGAGAGGCTCAGAAAGGAATCCA 34214

RESULT 5  
 AAV73804  
 ID AAV73804 standard; DNA; 35100 BP.  
 XX  
 AC AAV73804;  
 XX  
 DT 17-OCT-2003 (revised)  
 DT 25-FEB-1999 (first entry)  
 XX  
 DE KSHV LUR DNA (nucleotides 70,201-105,300).  
 XX  
 KW Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;  
 KW dihydrofolate reductase; LUR; long unique region; vaccine; prophylaxis;  
 KW diagnosis; treatment; HHV8; transactivator; LCTP; dUTPase; IEP-II;  
 KW DNA replication protein I; immediate early protein II; virF1; ICSBP;  
 KW phosphoprotein; DNA replication protein II; ribonucleotide reductase;  
 KW assembly/DNA maturation protein; tegument protein; ds.  
 XX  
 OS Human herpesvirus 8.  
 XX  
 PN US5849564-A.  
 XX  
 PD 15-DEC-1998.  
 XX  
 PF 29-NOV-1996; 96US-00770379.  
 XX  
 PR 29-NOV-1996; 96US-00770379.  
 XX  
 PA (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX  
 PI Bohenzky RA, Moore PS, Russo JJ, Chang Y, Edelman IS;  
 DR WPI; 1999-069741/06.  
 XX  
 XX Kaposi's sarcoma-associated herpes virus nucleic acid - encodes  
 PT dihydrofolate reductase and is useful for treatment, prophylaxis or  
 PT diagnosis of Kaposi's sarcoma.  
 XX  
 PS Disclosure; Col 125-156; 109pp; English.

XX This sequence is a fragment of the Kaposi's sarcoma-associated  
 CC herpesvirus (KSHV) LUR (long unique region). This fragment contains  
 CC coding regions for ORF48, ORF49, ORF50 which encodes a transactivator  
 CC (LCTP), K3, ORF52, ORF53, ORF54 which encodes dUTPase, ORF55, ORF56 which  
 CC encodes DNA replication protein I, ORF57 which encodes immediate early  
 CC protein II (IEP-II), K9 which encodes virF1 (ICSBP), K10, K11, ORF58  
 CC which encodes a phosphoprotein, ORF59 which encodes DNA replication  
 CC protein II, ORF60 which encodes a small ribonucleotide reductase, ORF61  
 CC which encodes a large ribonucleotide reductase, ORF62 which encodes an  
 CC assembly/DNA maturation protein, ORF63 which encodes tegument protein II,  
 CC ORF64 which encodes tegument protein III. KSHV is a new human Herpesvirus  
 CC (HHV8) believed to cause Kaposi's sarcoma (KS) which is the most common  
 CC form of neoplasia occurring in persons with acquired immune deficiency  
 CC syndrome (AIDS). The DHFR protein is useful for vaccination, prophylaxis,  
 CC diagnosis and treatment of a subject with Kaposi's sarcoma and for  
 CC detecting expression of a DNA virus associated with Kaposi's sarcoma in a  
 CC cell. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 35100 BP; 8632 A; 10010 C; 8650 G; 7808 T; 0 U; 0 Other;

Query Match 7.9%; Score 714; DB 2; Length 35100;  
 Best Local Similarity 54.4%; Pred. No. 8.6e-201;  
 Matches 1569; Conservative 0; Mismatches 1280; Indels 37; Gaps 5;  
 QY 5511 GGCACGGCTCATGCGCATTCACAGATTTCTTAAATCTGTTCCGAACATACGAGTCTAAAT 5570  
 Db 489 GACACTGGCCATGACATGCTCCACCGTCTCAGTCCGAGTAATGTGCTCTATGAGAT 548  
 QY 5571 CAAGTATACCAACAGCAAGTTCAACAGGAGGCAACGCGCTGGTTTCCACCGGACATC 5630  
 Db 549 GTGGTAGAGCAGCAGACGCTTCAACACGATGAATCAAGCTTAAGCTCCGCGGAACTC 608  
 QY 5631 CCGGCTCACAAACCCGAACTCTCTAGGAACGCACATCAAGCGTTTCAGACAAATTCGA 5690  
 Db 609 CACATCCACAAAGGGGTATTTGCTCGGTCTGTGTAATAGTCTGGAAATGAAATCTCAG 668  
 QY 5691 AACCGAAACCTTAGTTCTCAGCACCATCTGTTTAAATCTGCGGTAAAGAGACGAGCGC 5750  
 Db 669 AAAAGACACTGACCCCAAGGAGAACCTGGCGTCTTGCMAAGTTGATGAGCCCGCAGA 728  
 QY 5751 CACACACTGTCTGTGTGAGGATTCAGCTTAGTGTGCTGCGAGGCTTGTCTGACACGCGC 5810  
 Db 729 AAGATGTGTCTCCGTGGGACAAGAGCTTGGGGGGGAGAGATGGGCTACATGGGT 788  
 QY 5811 TAGCTCCCTGGCGGCTCGCGCACCTTTGACAGCCGACGCCAGTTCAGAGCGAGCA 5870  
 Db 789 GATTTCTTCTACACGCTCATACATTTGGTGGCACCACACAGGCGCTGTTCCAGTATCAGCAT 848  
 QY 5871 GAAGCGCTCTTTGCGAGTCCGCGCATATCTCGGCCATAGAGATTCGCTCGGTGATCTTT 5930  
 Db 849 AATCTATCTTTGAGTATCCACATCAAGTCAATGATGCTGTGTGCTGCGCATTT 908  
 QY 5931 TAGGTGCATATATCTGTGTGGGTAAACAACTCATCTCCCTGTAAAGGGGAATTAAGTC 5990  
 Db 909 TGCCCGCATGTACATTTCTCTCCACATATTTTAAACATCTGTATTAAGTAAAGTAGATT 968  
 QY 5991 CGTTCTTCTATCTGTTTCTCTGTAATCTATAGTTACATCTTTAAGTATAACCAAGACGC 6050  
 Db 969 CAGTCTGTGTGTGAGCCCCCCCCGGGGAAGCAGCGATATGCTTCAGGACCAACGAGGAGCC 1028  
 QY 6051 TAAACCCAGGATGGTCCAGACACCTTTTAAATAATCGCAAAAGCGACCGGATGGTCTATC 6110  
 Db 1029 TAAGAACCCCGGTGTCCGCGCTCCGGAACAGACCTCTGAGAATAACGCTCGGTCTTAC 1088  
 QY 6111 CGGTTTCCATGCTCTCAAGGTAGTCAACGATGATCTCCAGTTAGACTCTGTCTCCCT 6170  
 Db 1089 GAAACCCGATGTGTACCGATGCCAATCTGTGCGCTCCAGCTCTCACATTTTATC 1148  
 QY 6171 GTTACACCGTTAACTGGAAATGGAAACGACCATGACACCGTGAATCTCTGATGTCTCT 6230  
 Db 1149 TCCAAATACCGGAATTTGGGATACACACCTCCATGTTTCACTGATGATGATGATGATCTC 1208

Qy 6231 AAAAATAACCCGGAACAGAGCTAAATACCAATGACTGTACCCCTACCCCAAGCCAGC 6290  
 Db 1209 CCCACCCAAACCCATAGGACCCAGCTACAGCTTATCTCTCCCAATAATACCAAGGAGTGA 1268  
 Qy 6291 CCCCGTACTATT-----AGACCAGGGTGAAGTAAACCAAGCTATCTCT 6330  
 Db 1269 CCGCGGACTCATTAAGCCCGCCCGACGAACCAAGTACTGGTGGCAATGACACGCTCCCT 1328  
 Qy 6331 TTAATAAACCCATACGTTGGAGTTTGTAAAGTAAACAGCT-CGTGTATATTTTCAAGCCCTG 6389  
 Db 1329 TTAATAAAGTCAACCTTACTCCGCAAGGGGTAGTCTCTGTGTGAATATCTGTCAGGACAGC 1388  
 Qy 6390 TCATAAAATCGATACCGACGACCAATCAGGTAATTAACCTTTTATTTTCAAGATTCAGGA 6449  
 Db 1389 CACAAATAATGGCCAGATGACAAAGTAAAGATCGACC--TTTATTTGTATCTGAACA 1446  
 Qy 6450 GGGCGTGTTTACAGTGTGTAGTGGAGCATATCTGGCAATGGGAATGGCTGTATGAAT 6509  
 Db 1447 ATGCGTGTTTACAATGTGTAGTGGAGCAGAGTTTCGCCAAGCTCTACGTCGGAACAGT 1506  
 Qy 6510 CCACATTAAGTCTCGATCAGGTGCTCTGTGTACTTTTACTAGTGCAGATGAACCTTAAT 6569  
 Db 1507 CGGGTGTCAAGGCTCTTTATTAAGTGTTCGGTGTACTTGACCAAGCCGGAACTTAGGT 1566  
 Qy 6570 TTGGCGTGTTTAGTGTGTACCAACCAAAAAAATGCTGCGCTCGCTTTTAGAGCTCTA 6629  
 Db 1567 TGGGTCTGTACAGTGTGTACAGCAAAAAAAGATCGGCGGTGCTTTTTCAGGAGTGA 1626  
 Qy 6630 TGCAGCTTTAATCGTGTCAATAGCTTGTCTTTAAATGCAACTCTGTATCATCTTTA 6689  
 Db 1627 GGGACGTGTGATTATGTGACAGCTTCTGTCTGTAAATGCAAGCTGCTGTAATCTGAA 1686  
 Qy 6690 TGACAGTGTGTCCAAAAAACAAGATTTAAAAACACGTTAAATCGGTGCTCGTGTAGT 6749  
 Db 1687 CGACAGCTGTCCAAAAAACAAGTTCAGCTGCACGTTAAATCTGTATCTGAAAGT 1746  
 Qy 6750 CTTCTGTATTAATCTCTCAACAGAAAAATTTTAAACCAACTCGCCAGTACTGAA 6809  
 Db 1747 CTTCTGTAAATGACAGTTTCTACCAAGAAAAAATTTTACCGCTGGCCATCCACTGAA 1806  
 Qy 6810 ACAGTGTACGAGCAGGTGCTGAAAGTGTCTATCATCTCTCTCACTCCCTCTTTGAAGG 6869  
 Db 1807 AGGAGGAGCAGCTCCCGTTGTGCGTTGTAGGATATCCCTAACTTCGGAGCGGAGAC 1866  
 Qy 6870 TTTTGGTTACACGACCATGTGAGATAGGACACAGCTCAGGGGGAGTGGATCTGTAG 6929  
 Db 1867 GGGCGGACGCTCCCAAAAAATGGAGAGGACCACTCTGTGCACTCCGCGTCTGGGGTT 1926  
 Qy 6930 GGGGTAAATTCGTGGGCGTGGGGGTACTCCAGTTTGTAGATCTCCGCGAGACGCGTA 6989  
 Db 1927 CTGATTCAGGGCGCGGTGGGGGTATTTGGAGAGTCAAAACTCTCGGAGTCCCTTAA 1986  
 Qy 6990 CGAGTTCTTCGTCAAAATCTACGAGCCCGCATCCACAGGGGAGTACTGAGTTAATAA 7049  
 Db 1987 TGAGCTCTCTCAAAAACCTATGACGACGAGCTCCACTAGTGGCAGCATGCCCGTTAATA 2046  
 Qy 7050 GTCCCGAATTCATCTGTCTGTATTTTGTGAGCTCTTTTAGACATTAACCATGTTC 7109  
 Db 2047 CACCCCTTATCTGTGTGTGCTCAAGTTTGTACAACTGCTGAGGGAATTAAGCCAAATTC 2106  
 Qy 7110 TCCGACATCGGGTACGATGAACCGCTCTCTCTGTGTGTGATGACACACCTGGAGTA 7169  
 Db 2107 CCTAGCCCGGGAACAGGTACGGCTCGCTTTGTGTGTGTGAGCAATATCTGAATGG 2166  
 Qy 7170 TCTTAGACAGTATAGGCGCGTTCCTACTTAAGGGTTGCCGCTTGTGTAGAGCGGTTGC 7229  
 Db 2167 TCTTTGCAAGGTATAGGCTCTCTCAAGTTTGTAGCGGGTACGTTGGAGTCTGGATGA 2226  
 Qy 7230 GGGCATTCAGCTCAGTGTGTCTATTTCACCGCATATCATGCCAGTATCATGCCAGTGGCGGATAGTGA 7289  
 Db 2227 GGGTGGGACGAGGAGGTATCTAACTCTGAAATCTGATCCAGGACGGGTAAATGAT 2286  
 Qy 7290 ACTTGTACATATGATTTTAAAGATGGTCTTTTCCCGTAAATGGCACTCATGCTGTGTTTT 7349

Db 2287 ACCTAAACAGATGTTTGAACAGGTGATCTTTTAAAGGGGCTTCTCGATGTCTATGAAAAA 2346  
 Qy 7350 A-----ACGTTACAGGGTGTAAATAAACTCAGAGGTCTCTCGCTGGACACTTT 7398  
 Db 2347 CTATGACACGCACTCTCTCTTAGGGTAAAGAGCTTCGGCGGTCTCTGTGGAAAGCTT 2406  
 Qy 7399 TTGGGAGCTCTCCGAACCGGTACAGGATAAGCTGTACACCGCGGTCTCTTTATAAACA 7458  
 Db 2407 CGTGGGCTCTCGGACGAACCTGAAGGCCAACTTACCAAGTGTGTCTCTTATAAATGA 2466  
 Qy 7459 AGCATAGCTCAAAAAAATGACCCACGCTGACCTTAATAAGTCTACAGAGAGTCTCTCG 7518  
 Db 2467 CGCATACGAACAATCTACGATCCAGTGAACCTTAATAGAGTGGTGAAGATGTGTGAT 2526  
 Qy 7519 ATGTATTGTAAATGAAATGTGCTCCCTGGGTGCCATCTCCGGGCTCATCGCGGATCTGAA 7578  
 Db 2527 TCGATTTATGAAGAAATGTTTCCAAGCTTGTGTGCGCTATGTGCTGTCTTTACAGACATTA 2586  
 Qy 7579 TCTCTTTAACTGTGTCTCTACCGGGTTCGCGGTTTAAACCCCGCGGCGCGCAC 7638  
 Db 2587 CATGTTTAACTTGTCTCTTTTTCGTGCTCTCGAATGAGGACCAAGAGCGCGCGG 2646  
 Qy 7639 CTGTAAAGTCCCGTGGCGAGAAATCGCCAGGGCATCTGTGAGAAATCTCACCGAGAGGCG 7698  
 Db 2647 GTACAACTGTCATGCGCAGAGGCATCCCAAGGCATTTATTCGGATCTCTACGGAGAGAT 2706  
 Qy 7699 CCTTTGTTTACCAGAAAGATGTTTATAGCTGTCTCTGCGAGCGGGTCTGTATACCGCC 7758  
 Db 2707 CTTATTTGACAGAAAGGATTTCTGACAGCGCATGACAGCGGGTGAAGCTGCTCTCC 2766  
 Qy 7759 CCAAGTGGCCAGAGTCTGACAGAGCTCTACCGGAAATGAAGCCAAAGTGTTTGGGGGC 7818  
 Db 2767 AGCCATATGTAAGTACTACAGAAATATACACTGAAATGAAGCCAAATGCTTGGGGGC 2826  
 Qy 7819 GTGGCGTCTCTCATATGTTTTCAGCGCGCTTATCATGGCTATGTCGAGTCTCGTCTCTGT 7878  
 Db 2827 CTGGAGGCGACTGCTGCTCAATCGAGGCCCATATGATATTAACCTCTCTCCCTACTGAA 2886  
 Qy 7879 CACTTATTAATACCTGAGTGCAGGAGGAAACTGGAACCTTAGACTCTAAGACACTGTGAA 7938  
 Db 2887 GCTCTACAAACGTAACGATACCGCGGCTCTCTCTGAGCAGTCCAGGGCCCTCTGCT 2946  
 Qy 7939 ACTTCTTTTCAACCCATCTTTTCTTCAAGAAATCTTAGCCCTATGACGCTACTGGCAA 7998  
 Db 2947 TTTGTTTTCGAACCGGTCTACTTTCGAGGATTTATGCGCGCTGGAGATCATGACCAA 3006  
 Qy 7999 CGGGAAGATGTTCTGTGAAATTTTACCATCAACGCTAGCGGCGGAGAGAGCGGCC 8058  
 Db 3007 GGGTCAGCTCGCCCTGAAAACTTTTACAGCATCACCGGTTCTGCTGAGAAACGCGGCC 3066  
 Qy 8059 TGTCTGATGGAAGTACTAGCGGATGACGTGTCGGGAAAGCAGCTTGTCCCGATTC 8118  
 Db 3067 AATTACCCCGCAAGGTCACTGGACTGCTCTATCCAGGAAGCGGTCTCATGCCAGAAATC 3126  
 Qy 8119 CTTAATCTTCAGATGATGGAACCGGGCTTCTCCCGGCAACCCCTGTTGACCTCAGTAA 8178  
 Db 3127 TTTAATTTTGCATCTGGAGCCAGGACTGTTGCGGCTTCCATGTTAGACCTCAGGA 3186  
 Qy 8179 TGTCTTAGAATCCAGAAATCATCTCAGCGGCCCAACCCCTGAGTCAATTTGTCTATCAC 8238  
 Db 3187 TGTGTGGGAAACCGCGGTTTCTGAGCGCCCTGCTGAGCGGCTTGTCTATTAG 3246  
 Qy 8239 AAACACGCAACCCAGTCTGCTCTAGTCTAGTCAAGTATTAAGCCCAACCGAGGCGTGT 8298  
 Db 3247 CAAACCCATCCCAACATGCCGACACAGCTCAGCATCATCCCTTTAAACC---CATCGG 3303  
 Qy 8299 TCCGGCCATGTTTTTATGACACGTGGAAAGCGGTGTACAGAGCATTCACACAGAGC 8358  
 Db 3304 TACAGACCGCGGTTTATTAGTACGTGGAGCGGCCGCTCAGAAATATGTTGTACACAC 3363  
 Qy 8359 ACAGAC 8364

3364 ATCCAC 3369

RESULT 6

AAV62153/c

ID AAV62153 standard; DNA; 7361 BP.

AC AAV62153;

XX 18-DEC-1998 (first entry)

XX HSV-2 strain SB5 Contig ID 1 DNA sequence.

DE HSV-2 strain SB5; immunological response induction; therapy;

XX antiviral identification; viral protein inhibitor; ss.

XX Herpes simplex virus 2.

XX Key Location/Qualifiers

FT 322..864

FT CDS

FT /tag= a

FT /product= "ORF#1 protein"

FT /notes= "encoded protein shown in AAW72089"

FT 737..1741

FT CDS

FT /tag= b

FT /product= "ORF#2 protein"

FT /notes= "encoded protein shown in AAW72090"

FT 1796..2497

FT CDS

FT /tag= c

FT /product= "ORF#3 protein"

FT /notes= "encoded protein shown in AAW72091"

FT complement (2697..3302)

FT /tag= d

FT /product= "ORF#4 protein"

FT /notes= "encoded protein shown in AAW72092"

FT complement (3369..6014)

FT /tag= e

FT /product= "ORF#5 protein"

FT /notes= "encoded protein shown in AAW72093"

FT 5584..7361

FT CDS

FT /tag= f

FT /transl\_except= (pos: 7360..7361, aa: Arg-Val)

FT /product= "ORF#6 protein"

FT /notes= "encoded protein shown in AAW72094"

XX WO9820016-A1.

XX 14-MAY-1998.

XX 31-OCT-1997; 97WO-US020016.

XX 04-NOV-1996; 96US-0030279P.

XX 09-JUN-1997; 97US-0049018P.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

XX Esser KM, Chan JY, Dabrowski-Amaral CE, Delvecchio AM, Dillon SB;

XX Leary JJ;

XX WPI; 1998-286847/25.

XX P-PSDB; AAW72089, AAW72090, AAW72091, AAW72092, AAW72093, AAW72094.

XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention and

XX treatment of infection or inducing immunological response in mammal.

XX Claim 1; Page 338-341; 748pp; English.

XX This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA

XX sequence of the invention. This sequence was isolated from HSV-2 strain

XX SB5 (deposited as ATCC VR-2546), is designated contig ID 1, and encodes 6

XX HSV-2 proteins. The proteins can be used for the treatment or prevention

XX of disease, to induce an immunological response in a mammal or to

XX identify inhibitors, activators or novel antivirals. Antagonists of the



```
RESULT 7
AAV62176/c
ID  AAV62176 standard; DNA; 117213 BP.
XX  AC  AAV62176;
XX  DT  13-JAN-1999 (first entry)
XX  DE  HSV-2 strain SB5 Contig ID 15 DNA sequence.
XX  KW  HSV-2 strain SB5; immunological response induction; therapy;
XX  KW  antiviral identification; viral protein inhibitor; ss.
XX  OS  Herpes simplex virus 2.
XX  FH  Key Location/Qualifiers
XX  CDS  755..1297
XX  FT  /tag= a
XX  FT  /product= "ORF#1 protein"
XX  FT  /note= "encoded protein shown in AAW72170"
XX  CDS  1170..2174
XX  FT  /tag= b
XX  FT  /product= "ORF#2 protein"
XX  FT  /note= "encoded protein shown in AAW72171"
XX  CDS  2229..2930
XX  FT  /tag= c
XX  FT  /product= "ORF#3 protein"
XX  FT  /note= "encoded protein shown in AAW72172"
XX  CDS  complement(3130..3735)
XX  FT  /tag= d
XX  FT  /product= "ORF#4 protein"
XX  FT  /note= "encoded protein shown in AAW72173"
XX  CDS  complement(3802..6447)
XX  FT  /tag= e
XX  FT  /product= "ORF#5 protein"
XX  FT  /note= "encoded protein shown in AAW72174"
XX  CDS  6017..8482
XX  FT  /tag= f
XX  FT  /product= "ORF#6g protein"
XX  FT  /note= "encoded protein shown in AAW72250"
XX  CDS  6026..8482
XX  FT  /tag= g
XX  FT  /product= "ORF#6f protein"
XX  FT  /note= "encoded protein shown in AAW72249"
XX  CDS  6085..8482
XX  FT  /tag= h
XX  FT  /product= "ORF#6e protein"
XX  FT  /note= "encoded protein shown in AAW72248"
XX  CDS  6167..8482
XX  FT  /tag= i
XX  FT  /product= "ORF#6d protein"
XX  FT  /note= "encoded protein shown in AAW72247"
XX  CDS  6296..8482
XX  FT  /tag= j
XX  FT  /product= "ORF#6c protein"
XX  FT  /note= "encoded protein shown in AAW72246"
XX  CDS  6326..8482
XX  FT  /tag= k
XX  FT  /product= "ORF#6b protein"
XX  FT  /note= "encoded protein shown in AAW72245"
XX  CDS  6446..8482
XX  FT  /tag= l
XX  FT  /product= "ORF#6a protein"
XX  FT  /transl_except= (pos: 7400..7402, aa: Ala-Ala)
XX  FT  /transl_except= (pos: 7481..7486, aa: Ile)
XX  FT  /note= "encoded protein shown in AAW72244"
XX  CDS  8457..9347
XX  FT  /tag= m
XX  FT  /product= "ORF#7 protein"
XX  FT  /note= "encoded protein shown in AAW72175"
XX  CDS  complement(9604..11855)
XX  FT  /tag= n
XX  FT  /transl_except= (pos: 11635..11636, aa: Ala)
XX  FT  /product= "ORF#8 protein"
XX  FT  /note= "encoded protein shown in AAW72176"
XX  CDS  complement(11905..14520)
XX  FT  /tag= p
XX  FT  /product= "ORF#9a protein"
XX  FT  /note= "encoded protein shown in AAW72223"
XX  CDS  complement(11905..14508)
XX  FT  /tag= o
XX  FT  /product= "ORF#9b protein"
XX  FT  /note= "encoded protein shown in AAW72222"
XX  CDS  14399..15802
XX  FT  /tag= q
XX  FT  /product= "ORF#10 protein"
XX  FT  /note= "encoded protein shown in AAW72177"
XX  CDS  complement(15996..16286)
XX  FT  /tag= x
XX  FT  /product= "ORF#11 protein"
XX  FT  /note= "encoded protein shown in AAW72178"
XX  CDS  complement(16202..18064)
XX  FT  /tag= s
XX  FT  /product= "ORF#12 protein"
XX  FT  /note= "encoded protein shown in AAW72179"
XX  CDS  complement(18105..19661)
XX  FT  /tag= t
XX  FT  /product= "ORF#13 protein"
XX  FT  /note= "encoded protein shown in AAW72180"
XX  CDS  complement(19415..20074)
XX  FT  /tag= u
XX  FT  /product= "ORF#14 protein"
XX  FT  /note= "encoded protein shown in AAW72181"
XX  CDS  20155..21453
XX  FT  /tag= v
XX  FT  /product= "ORF#15 protein"
XX  FT  /note= "encoded protein shown in AAW72182"
XX  CDS  complement(21326..22291)
XX  FT  /tag= w
XX  FT  /product= "ORF#16 protein"
XX  FT  /note= "encoded protein shown in AAW72183"
XX  CDS  complement(22546..24654)
XX  FT  /tag= x
XX  FT  /product= "ORF#17 protein"
XX  FT  /note= "encoded protein shown in AAW72184"
XX  CDS  24684..25955
XX  FT  /tag= y
XX  FT  /product= "ORF#18 protein"
XX  FT  /note= "encoded protein shown in AAW72185"
XX  CDS  complement(26295..27251)
XX  FT  /tag= z
XX  FT  /product= "ORF#19 protein"
XX  FT  /note= "encoded protein shown in AAW72186"
XX  CDS  complement(27630..31784)
XX  FT  /tag= ab
XX  FT  /product= "ORF#20a protein"
XX  FT  /note= "encoded protein shown in AAW72224"
XX  CDS  complement(27630..31754)
XX  FT  /tag= aa
XX  FT  /product= "ORF#20b protein"
XX  FT  /note= "encoded protein shown in AAW72225"
XX  CDS  complement(32067..32735)
XX  FT  /tag= ac
XX  FT  /product= "ORF#21 protein"
XX  FT  /note= "encoded protein shown in AAW72187"
XX  CDS  33140..34984
XX  FT  /tag= ad
XX  FT  /product= "ORF#22a protein"
XX  FT  /note= "encoded protein shown in AAW72226"
XX  CDS  33386..34984
XX  FT  /tag= ae
XX  FT  /product= "ORF#22b protein"
XX  FT  /note= "encoded protein shown in AAW72227"
XX  CDS  complement(35205..37721)
```

```
FT      /*tag= af
FT      /product= "ORF#23 protein"
FT      /note= "encoded protein shown in AAW72188"
FT      complement(38058. .39188)
FT      CDS
FT      /*tag= ag
FT      /product= "ORF#24 protein"
FT      /note= "encoded protein shown in AAW72189"
FT      39090. .39935
FT      CDS
FT      /*tag= ah
FT      /product= "ORF#25 protein"
FT      /note= "encoded protein shown in AAW72190"
FT      40216. .41973
FT      CDS
FT      /*tag= ai
FT      /product= "ORF#26 protein"
FT      /note= "encoded protein shown in AAW72191"
FT      42206. .44178
FT      CDS
FT      /*tag= aj
FT      /transl_except= (pos: 44063. .44064, aa: Lys)
FT      /product= "ORF#27 protein"
FT      /note= "encoded protein shown in AAW72192"
FT      complement(44853. .47297)
FT      CDS
FT      /*tag= ak
FT      /product= "ORF#28 protein"
FT      /note= "encoded protein shown in AAW72193"
FT      47122. .47338
FT      CDS
FT      /*tag= al
FT      /product= "ORF#29 protein"
FT      /note= "encoded protein shown in AAW72194"
FT      complement(47305. .49662)
FT      CDS
FT      /*tag= am
FT      /product= "ORF#30 protein"
FT      /note= "encoded protein shown in AAW72195"
FT      complement(50035. .51666)
FT      CDS
FT      /*tag= an
FT      /product= "ORF#31 protein"
FT      /note= "encoded protein shown in AAW72196"
FT      complement(51701. .53575)
FT      CDS
FT      /*tag= ao
FT      /product= "ORF#32 protein"
FT      /note= "encoded protein shown in AAW72197"
FT      54393. .58115
FT      CDS
FT      /*tag= ap
FT      /product= "ORF#33 protein"
FT      /note= "encoded protein shown in AAW72198"
FT      complement(58060. .58977)
FT      CDS
FT      /*tag= aq
FT      /product= "ORF#34 protein"
FT      /note= "encoded protein shown in AAW72199"
FT      complement(58970. .60760)
FT      CDS
FT      /*tag= ar
FT      /product= "ORF#35 protein"
FT      /note= "encoded protein shown in AAW72200"
FT      60759. .61151
FT      CDS
FT      /*tag= as
FT      /product= "ORF#36 protein"
FT      /note= "encoded protein shown in AAW72201"
FT      61241. .62071
FT      CDS
FT      /*tag= at
```

```
Query Match      2.2%; Score 200.8; DB 2; Length 117213;
Best Local Similarity 52.3%; Pred. No. 8.4e-47;
Matches 495; Conservative 0; Mismatches 442; Indels 9; Gaps 2;

QY 585 CAATATATTGTAATCGACGAGCGCGAACCCGTCTCTTACATACCTACCCACCGTCTG 644
DB 5722 CAACATTATCGTATCGACGAGCGCGGCTCTTGGGGCGGCACCTACTCAGACCGTGT 5663
QY 645 GTTCTTTTACTGGTTCCTAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGCGGT 704
DB 5662 GTATTGCTGGTATGATTAAACCCCTGTACCACACCCCGGACGTACGCGCGCGCTGG 5603
QY 705 TCCGTGATAGTATGCTCGGTTCGCCAAGCAGAGAAACGGCTTCAGTCTCACTACAA 764
```

```
DB 5602 GCCGGTGTGCTGTGCTGCGGTCGCCGACCCAGACGGCCTCGCTGGAGTCCACCTTCGA 5543
QY 765 CCACGGGACGCAAAAGACGAGATATCGTGTGGAGAAATCATCTACATTCATGATAGG 824
DB 5542 ACACCAAAACTCGGATGCTCCGTCGCGAGAGAAAGCTGTCTACGTACCTCTCTG 5483
QY 825 GAAGAAGGTGCTATCTCTGAGTACGTACACTTGGAGAGAAACTGGGCGCTGTTTATAA 884
DB 5482 CAACCGCACCTACGCGAGTACACGCGCTCTCGACAGCTGGCCATTTTCATTAACAA 5423
QY 885 CAACCGCTGCACCGATCTGCAGTTCGACACCTTCTAAATAATTTAGAGTATATCTTCC 944
DB 5422 CAACGCGATGTGTGAGCACGAGTTCGGGAACCTCATGAAGTGTGTGAGTACGGCC 5363
QY 945 CATTCCTGACGAAGTCTAGTTCACGTAGACAGATTTGTCTCTCTCTCTCTCTCTCTCT 1004
DB 5362 CATCACCGAGGAGCACATGCGATTTGTGGACCGCTTTGTCTCTCTCTCTCTCTCTCT 5303
QY 1005 GGATCTTTTAGATACATTTGGCTGGACCGCACTCTTTTGTCTACATAGCAGGTAAAG 1064
DB 5302 CAACCGCGCAACCTTCCGGGCTGGACGCGCTTCTCTCTCTCTCTCTCTCTCTCTCT 5243
QY 1065 GTATCTAACAAACCTACACACATGTCTAAGCTAGGGGGGCGATACACGAGACACGAAG 1124
DB 5242 GTACATGCCCAAGCTCCACGCTTACCTAAGGTGACTCGCA---GGGGAGTTTGTGT 5186
QY 1125 CTTTACCTGTCCCGTGTGTGCGAGGTGTTTGTGAACCGTTCGAGGAATACAGCGGCG 1184
DB 5185 GTTTACCTGTCCCGTGTGTGCGGTTCGAGGTTCGAGGTATCGACGCT 5126
QY 1185 CGTCAACTCACCAACCTCACGCGTACCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1244
DB 5125 CACGAGCAACCCACGCTGACCATGGAAGTGGATCAGGCGCAACGCCAGTCCGATCAC 5066
QY 1245 TAATTATTGCGAGTTTGTGGACGAGCATGTCCATAGTCGCCACGGAATC-----CAC 1298
DB 5065 CACTTACTCCAGAGTCAGGACACGAGCGCGGCGCGTGTGAGGTGCGACAGCAA 5006
QY 1299 TGAAGTTCGACCCAGGTACCTTTATCCCAAGTTGTTTAAACAGCCACGCGTCTTT 1358
DB 5005 GCAACAGCTAGTGTGCGCGGGAAGACATCAGCTACTCTCTCAACAGCCAGGTTCGG 4946
QY 1359 AAATGGAAAAACAAAAAATGCATATATGTGGTTCAGGGTACGCTACTTTGAGTTCAA 1418
DB 4945 GACCGCGCGCTCCGAAAGATGGTGTGTTGGTTTCGACGGGACGTTTCGACCTTCG 4886
QY 1419 AATCTTAGACAGCGAATCTTCTGTGGAGACGCAATCCGACAGCCGTCCGGAATAGT 1478
DB 4885 TGTCTGCGGACGACAGCTTCGTGAAGACCCAGGGGAGACCTCGGTGGAGTTTCGC 4826
QY 1479 CGGTTTCTTAAACACGCTACTGTACAAACGCGCATCTACTCTCTCTCTCTCTCT 1524
DB 4825 CCGTTTCTGTGCGGCTCATGTTCCGGGGGCTGATTCACCTTTTAC 4780

RESULT 8
AAD25519/c
ID AAD25519 standard; DNA; 154746 BP.
XX
AC AAD25519;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human herpesvirus 2 complete DNA genome.
XX
KW Human herpesvirus 2; cytotstatic; cancer; immunosuppressive; virucide;
KW antibacterial; fungicide; protozoacide; antirheumatic; antiinflammatory;
KW antiarthritic; rheumatoid arthritis; neuroprotective; multiple sclerosis;
KW immune response; vasotropic; vaccine; gene therapy; autoimmune disease;
KW vasculitis; ds.
XX
OS Human herpesvirus 2.
XX
```

PN WO200176643-A1.  
 XX 18-OCT-2001..  
 PD 06-APR-2001; 2001WO-US011372.  
 XX 07-APR-2000; 2000US-0195680P.  
 PF (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX Orson FM, Kinsey BM, Bhogal BS;  
 XX WPI; 2002-066308/09.  
 XX  
 XX Composition for oral delivery of vaccines, comprises expression vector  
 XX containing antigenic genomic sequence, bound to aggregated protein-  
 XX polycationic polymer conjugate or suspension.  
 XX  
 XX Disclosure; Page 90-132; 145pp; English.  
 XX  
 XX The invention relates to a composition comprising an expression vector  
 XX bound to an aggregated protein-polycationic polymer conjugate or  
 XX suspension. The expression vector contains a promoter polynucleotide  
 XX sequence operatively linked to a polynucleotide sequence encoding an  
 XX antigen which is a fragment of a gene or genome associated with an  
 XX infectious disease, cancer and autoimmune disease such as rheumatoid  
 XX arthritis, vasculitis, and multiple sclerosis, pathogenic genomes  
 XX consisting of bacterium, fungus, protozoa and virus such as human  
 XX immunodeficiency virus (HIV) herpes simplex virus (HSV), hepatitis C  
 XX virus (HCV), influenza and respiratory syncytial virus (RSV), and  
 XX optionally comprising a nucleotide sequence encoding a cytokine (or a  
 XX cytokine expression vector), is useful for inducing an immune response  
 XX (systemic and/or mucosal) in an organism. The cytokine expression vector  
 XX contains a sequence for granulocyte macrophage-colony stimulating factor  
 XX (GM-CSF) and interleukin-12 (IL-12). The polynucleotide sequences encoding  
 XX the antigen and the cytokine are under transcriptional control of same or  
 XX different promoter polynucleotide sequences. The expression vector, as a  
 XX DNA vaccine is useful for treating a condition in an organism. The  
 XX present sequence is human herpesvirus 2 complete DNA genome related to  
 XX the invention  
 XX  
 XX Sequence 154746 BP; 23003 A; 54218 C; 54701 G; 22824 T; 0 U; 0 Other;  
 XX  
 XX Query Match 2.2%; Score 200.8; DB 6; Length 154746;  
 XX Best Local Similarity 52.3%; Pred. No. le-46;  
 XX Matches 495; Conservative 0; Mismatches 442; Indels 9; Gaps 2;  
 XX  
 XX 585 CAATATTATTGTAATCGACGAGGCGGAAACCTCTCTTACATACCTACCCCGTCGT 644  
 XX 14524 CAACATATCTGTCATCGACGAGGCGGGCTCTCGGGCGGCACCTACTCAGACCGTGGT 14465  
 XX 645 GTTCTTTTACTGTTCTTAACACGTGGCTTAATACCCCTCTTTTACCGCCAGCGGGCGGT 704  
 XX 14464 GTATTGCTGGTGATGATTACGCCCTGTACACACCCCCCAGTACCGCGGCGCGCTGCG 14405  
 XX 705 TCCGTGCATAGTATGCGTGGTTCGCCAAGCAGACAAACCGCTTCCAGTCAACTTACAA 764  
 XX 14404 GCCGTGCTGTTGCTGGGGTTCGCCAAGCAGACCGCTTCCAGTCCACCTTCA 14345  
 XX 765 CCACGGGACGCAAAAGACGAGATATGCTGTCGAGAAATCTTAATCATGATAGG 824  
 XX 14344 ACACCAAAATGCGATCTCGTCCGCGCAGAGCGAAACGTCTCAGTACCTCATCTG 14285  
 XX 825 GAAGAAGTCTATCTGAGTACGTACCTTGGAGAGAACTGGCGGCTTTTATAACAA 884  
 XX 14284 CAACCGACCTTACCGAGTACCGGCTCTCGACAGCTGGGCCATTTTCAATACAA 14225  
 XX 885 CAAGCGCTGCACCGATTCGAGTTTCGGACACTTCTCTAAATTTTAGATATATCTTCC 944  
 XX 14224 CAAGCGATGTGGGAGCAGAGTTCGGAACTCATGAAGGTGCTGGAGTACGCGCTTCC 14165  
 XX 945 CATTCCTGACGAGTTCATGAGTACGAGATTTGCTGCTTCTTAAAGTAAGATTAT 1004

Db 14164 CATCACCGAGGAGCAGCATGTCAGTCTTGTGGACCGCTTTGTCTCCGGAAGTTTACATCAC 14105  
 Qy 1005 GGATCCCTTTAGATACATTTGGCTGGACCCGACTCTTTTGTGTCATAGCGAGGTAAAGGC 1064  
 Db 14104 CAACCGGCGCAACCTTCGGGGTGGACGGCTGTTCTGCTCCACAGAGGTCGACGGC 14045  
 Qy 1065 GTATCTAACAAACCTACACATCTTAACGCTAGGGGGCGATACACGAGGACACGAAGCT 1124  
 Db 14044 GTACATGGCCAAAGCTCCAGGCTACCTAAAGGTGACTCGCGA---GGGGGAGTTTGTGT 13988  
 Qy 1125 CTTTACCTGTCCTGGTGTGGAGGTGTTGTGAAGCCGTTTTCAGGAATACAAAGCGGC 1184  
 Db 13987 GTTTACCTCTCCCTGCTGCTCAGCTTTGTGTCGTTCAAGAGTTTGAAGATATCAGCGCT 13928  
 Qy 1185 CGTCAACCTCACCAACCTCACCGTGACCGAGTGGGTGACAAAAAACCCTCTTTAAAGTTAAG 1244  
 Db 13927 CAGCGAGCAACCCAGCTGACCATGGAAAAGTGATCAGGCCAACGCCAGTCGCATCAC 13868  
 Qy 1245 TAATATTTCGAGTGTGGGACCGAGCATGTCATAGTCCCGACGGAATC-----CAC 1298  
 Db 13867 CAATCTATCCAGAGTCAGGACCGGCGGCGACGTCGCTGTGAGGTGACAGCA 13808  
 Qy 1299 TGAACGTTTCGACCCAGGTTTACCTTTATCACCAGTGTGTTTAAACACGCCAGCTGTCTTT 1358  
 Db 13807 GCAACAGTAGTGTGTCGGCCGGAACGACATCACGTACGTCTCAACGCCAGGTGCGGT 13748  
 Qy 1359 AAATCGAAAAACAAAAATGATATGCGGTTTCAGGTTACGTACTTTGAGTTCAAAAG 1418  
 Db 13747 GACCGCGCGCTCCGAAAGATGGTGTGGGTTCGACGGGACGTTTCGACCTTCGAGGC 13698  
 Qy 1419 AATCTAGACGAGCACTCTTCGTGGACGACCATTCGCAAGACCGCTCCGGAATAGTATA 1478  
 Db 13687 TGTGTCGGGACGACACTTCGTGAAGACCCAGGGGAGACCTCGGTGGAGTTGCGCTA 13628  
 Qy 1479 CGGTTTCTTAAACACGCTACTGTATCAACGCCCATGTACTGTTTTCAC 1524  
 Db 13627 CCGGTTCTGTCGCGCTCATGTTCCGGCGGCTGATTCACTTTTAC 13582  
 XX  
 XX RESULT 9  
 XX AAV62130  
 XX ID AAV62130 standard; DNA; 8952 BP.  
 XX AC AAV62130;  
 XX DT 07-DEC-1998 (first entry)  
 XX DE HSV-2 strain SB5 Contig ID 100 DNA sequence.  
 XX KW HSV-2 strain SB5; immunological response induction; therapy;  
 XX KW antiviral identification; viral protein inhibitor; ss.  
 XX OS Herpes simplex virus 2.  
 XX FH Location/Qualifiers  
 XX CDS complement(1..1802)  
 XX /tag= a  
 XX /product= "ORF#1 protein"  
 XX /transl\_except= (pos: 1..2, aa: Ala-Xaa)  
 XX /note= "Xaa= unspecified amino acid"  
 XX CDS 1372..4017  
 XX /tag= b  
 XX /product= "ORF#2 protein"  
 XX 4084..4689  
 XX /tag= c  
 XX /product= "ORF#3 protein"  
 XX complement(4889..5407)  
 XX /tag= d  
 XX /product= "ORF#4 protein"  
 XX complement(5646..6650)  
 XX /tag= e  
 XX /product= "ORF#5 protein"  
 XX complement(6523..7074)  
 XX CDS

FT FT /\*tag= f  
PN /product= "ORF#6 protein"  
XX WO9820016-Al.  
XX 14-MAY-1998.  
XX 31-OCT-1997; 97WO-US020016.  
XX 04-NOV-1996; 96US-0030279P.  
PR 09-JUN-1997; 97US-0049018P.  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX Esser XM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB,  
PI Leary JJ;  
XX WPI; 1998-286947/25.  
DR P-P8DB; AAW72001, AAW72002, AAW72003, AAW72004, AAW72005, AAW72006.  
XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention and  
PT treatment of infection or inducing immunological response in mammal.  
XX Claim 1; Page 163-166; 749pp; English.  
XX This sequence represents a Herpes simplex virus type-2 (HSV-2) DNA  
CC sequence of the invention. This sequence was isolated from HSV-2 strain  
CC SB5 (deposited as ATCC VR-2546), is designated Contig ID 100, and encodes  
CC 6 HSV-2 proteins. The proteins can be used for the treatment or  
CC prevention of disease, to induce an immunological response in a mammal or  
CC to identify inhibitors, activators or novel antivirals. Antagonists of  
CC the proteins can be used to inhibit a viral polypeptide. The DNA sequence  
CC or a vector containing it can also be used to induce an immunological  
CC response in a mammal  
XX  
XX Query Match 2.2%; Score 199.2; DB 2; Length 8952;  
XX Best Local Similarity 52.2%; Pred. No. 3.8e-47;  
XX Matches 494; Conservative 0; Mismatches 443; Indels 9; Gaps 2;  
QY 585 CAATATTATTCGACGAGGCGGACCCCTGCTCTTACATATCTACACCGCTGT 644  
DB 2097 CAACATTATGCTATCGACGAGGCGGCTCTGGGCGGCACCTACTACGACCGTGT 2156  
QY 645 GTCTTTTACTGGTCTTAACACGCTGGTAATACCCCTCTTTACCGCCAGGGGGGT 704  
DB 2157 GTATTGCTGGTGAATGATTAACGCTGTACCAACCCCGGAGTACGCGGCGCGCTCG 2216  
QY 705 TCCGTGCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 764  
DB 2217 GCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2276  
QY 765 CCACGGGACGAAAGACGAGATATGCTGCTGCGAGACATCTTAACATTCATGATAG 824  
DB 2277 ACACCAAACTGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2336  
QY 825 GAAGAAGTCTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 884  
DB 2337 CAACCGCACCTAGCGAGTACACGCGCTCTGCGACAGTGGGCGCATTTTCAATACAA 2396  
QY 885 CAGCGCTGCACCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944  
DB 2397 CAGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2456  
QY 945 CATTCCTGACGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004  
DB 2457 CATCACGAGGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2516  
QY 1005 GATCTCTTAGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064  
DB 2517 CAACCCGCGCAACCTTCCGCGGTGGACGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2576

QY 1065 GTATCTAACAAACCTACACACATGTCTTAACGCTAGGGGGCGATACGAGGACACGAAGCT 1124  
DB 2577 GTACATGGCCAAAGCTCCACGCTACCTAAAGTGACTCGCGA---GGGGAGTTTCTTGT 2633  
QY 1125 CTTTACCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1184  
DB 2634 GTTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2693  
QY 1185 CGTCAACCTCACCACCTCAGGAGTGGGTGACAAAAAAGCTCTTTAAGTTAAG 1244  
DB 2694 CACGACGAAACCCACGCTGACCATGGAAGAGTGGATCAGCGCCACGCGCATCAC 2753  
QY 1245 TAATTTTCGCGTTCGCGACGAGACATGTCCATAGTCGCGCACGGAATC-----CAC 1298  
DB 2754 CAATCTATCCAGAGTCCAGACCCAGGACGCGGGGCGCTGCTGCTGCTGCTGCTGCTGCT 2813  
QY 1299 TGAACGTTTCGACCGAGTTTACCTTTATCACCAGTTTGTAAAGAACAGCCACGCTGCTTT 1358  
DB 2814 GCAACAGCTAGTCTGTCGCCCGGAACGACATCACGTACGTCTCTCAACAGCGCGCTCGCGT 2873  
QY 1359 AAATGGAAAAAACAATAATGCTATGTGGTTTTCAGGGTACGTACTTTGAGTTCAAAAG 1418  
DB 2874 GACCGCGGCTCCGAAGATGCTGTTGGTTTCGCGGACGTTTCGCGACCTTCGAGGC 2933  
QY 1419 AATCCTAGACAGGAACTTTCGAGAGCGCATTCGAGACCGTCCGGAATACGTATA 1478  
DB 2934 TGTCTCGCGGACGACAGCTTCGTGAAGACCCAGGGGAGACCTCGGTGAGTTCGCTTA 2993  
QY 1479 CGGTTTCTTAAACACGCTACTGTACAAACGCGCATGTACTGTTTTCAC 1524  
DB 2994 CCGGTTCTGTCGCGGCTCATGTTCCGCGGCTGATTCATTITTAC 3039

RESULT 10

AAH74202

ID AAH74202 standard; DNA; 125157 BP.

XX AC AAH74202;

XX DT 15-OCT-2001 (first entry)

XX DE Nucleotide sequence of the Varicella virus Oka strain.

XX KW Oka strain; pox vaccine; vaccine; poxvirus disease; chickenpox; ss.

XX OS Varicella virus.

XX PN WO200156600-Al.

XX PD 09-AUG-2001.

XX PF 31-JAN-2001; 2001WO-JP000678.

XX PR 31-JAN-2000; 2000JP-00062734.

XX PA (OSAU ) UNIV OSAKA.

XX PI Gomi Y, Sunamachi H, Takahashi M, Yamanishi K;

XX XX WPI; 2001-488845/53.

XX PT Method for checking absence of mutation at specific positions of

XX PT varicella virus genome for quality control of attenuated live varicella

XX PS Disclosure; Page 106-150; 159pp; Japanese.

XX CC The present sequence represents the nucleotide sequence of Varicella

XX CC virus, Oka strain. The specification describes a method for the quality

XX CC control of an attenuated pox vaccine, characterised in that the DNA

XX CC sequence of specific parts of the viral genome in a sample is determined

XX CC and proved to be conserved rather than mutated in comparison with the

XX CC same parts of a reference viral genome. The method is useful for quality

|           |   |
|-----------|---|
| CC        | control of vaccines for use in the prevention of poxvirus diseases such     |
| CC        | as chickenpox   |
| XX        |   |
| SQ        | Sequence 125157 BP; 33776 A; 29439 C; 28285 G; 33624 T; 0 U; 33 Other;      |
|           | Query Match 1.7%; Score 154; DB 5; Length 125157;                           |
|           | Best Local Similarity 56.2%; Pred. No. 9e-33;                               |
|           | Matches 289; Conservative 0; Mismatches 225; Indels 0; Gaps 0;              |
| QY        | 585 CATATATTGTGAATCGAGGCGCGAACCCTGTCCTTTACACTCATCAACACGGTGT 644             |
| Dd        | 96652 CAACGTTATTGTTATTGATGAAGCAGGATGTAGGGCGTCATATCTCTCGGCCGTTGT 96711       |
| QY        | 645 GTTCTTTTACTGGTTCTCTAAACAGCTGGCTAAATACCCCCTTTTACGCCAGGGCGGT 704          |
| Dd        | 96712 TTACTCTTGTTGSCITTTGAATGCATATATCAAAGCCCTCAGTACATAAACGGTCGAA 96771      |
| QY        | 705 TCCGTCGATAGTAGTCGTGCGTTTCGCCAACGGCAGACAACGGGTTCAGTCAACTACAA 764         |
| Dd        | 96772 ACCGGTCATAGTAGTCGTGCGTTTCGCCCAACCCCAACCTGACTCGTTAGAATCTCATTTTCA 96831 |
| QY        | 765 CCACGGGAGCGCAAAAACGAGAGATATCGTCTGGCAGAACATCCCTTAAACATTCATGATAGG 824     |
| Dd        | 96832 ACATGACATCGACGGCTTCACAGTAACTCCCTAGTCAAAATATACTCACGTATATACTG 96891     |
| QY        | 825 GAAGAAGTCGTATCTGAAGTAGTACACTTGGAGAGAAACTGGGGCTGTTTTATAACAA 884          |
| Dd        | 96892 CAACTAAATCTGCGTCATATACTAACTCTCACATACCTGGGCAATCTTTTATTAATAA 96951      |
| QY        | 885 CAAGCGCTGCACCGATCTCGAGTCGGACACCTCTCTAAAAATTTTAGAGTATAATCTTCC 944        |
| Dd        | 96952 CAAACGATCTCAAGAGGACGATTTTGGAAATCTTTTAAACGCTTGAGTACGGGCTACC 97011      |
| QY        | 945 CATTCTGACGAAGTCATGATTAGTAGACAGATTTCTGTTCTTAAAGTAAGATTAT 1004            |
| Dd        | 97012 TATTACCGAAGCACATGCGCGTCTGGTCGATACATTTGTTACCTGCATCCTATATTA 97071       |
| QY        | 1005 GGATCCCTTTAGAAATACATTTGGCTGGACCCGACTCTTTTGTCAATACGAGGTAAAGGC 1064      |
| Dd        | 97072 CAACTCTGCTAATCTCCCGGATGACGCGTCTGTATTCTGTCGCAATAGGAGGTGCGC 97131       |
| QY        | 1065 GTATCTTAACAAACCTACACATGCTTAACGCTA 1098                                 |
| Dd        | 97132 GTATATGATAGTTACACGCGCATTTAAACTA 97165                                 |
| RESULT 11 |   |
| AAT85076  |   |
| ID        | AAT85076 standard; DNA; 2871 BP.  |
| XX        |   |
| AC        | AAT85076;   |
| XX        |   |
| DT        | 18-MAR-1998 (first entry)   |
| XX        |   |
| DE        | Human cytomegalovirus (HCMV) UL105 open reading frame.                      |
| XX        |   |
| KW        | Human cytomegalovirus primase; HCMV UL105; open reading frame; orf;         |
| KW        | screening; inhibitor; infection; diagnosis; ss.                             |
| XX        |   |
| OS        | Homo sapiens.   |
| XX        |   |
| Key       | Location/Qualifiers   |
| FF        | 1..2871   |
| CDS       | /tag= a   |
| FT        | /product= "HCMV_UL105_orf_translation_product"                              |
| FT        |   |
| FT        |   |
| PN        | GB2311069-A..   |
| XX        |   |
| PD        | 17-SEP-1997.  |
| XX        |   |
| PF        | 05-MAR-1997; 97GB-00004577.   |
| XX        |   |
| PR        | 14-MAR-1996; 96US-0013546P.   |

|    |  |
|----|--|
| PR | 04-APR-1996; 96GB-00007117.  |
| XX | (MERI ) MERCK & CO INC.  |
| PA |  |
| XX |  |
| PI | Gotlib L, Hazuda DJ, Lafemina RL;  |
| XX |  |
| DR | WFI; 1997-427907/40.   |
| DR | P-FSDB; AAW27086.  |
| XX |  |
| PT | Novel human cytomegalovirus primase - used for screening for HCMV          |
| PT | antivirals and in diagnosis of HCMV related diseases.                      |
| XX |  |
| PS | Example 1; Page 20-22; 28pp; English.                                      |
| XX |  |
| CC | The present sequence represents the human cytomegalovirus (HCMV) UL105     |
| CC | open reading frame (orf) spanning nucleotides 151926 through 154793 of     |
| CC | HCMV AD169 which was amplified by PCR primers AI185074-5. The amplified    |
| CC | product was digested with BglII and EcoRI and cloned into BgIII/EcoRI      |
| CC | digested pB87 vector DNA to yield the plasmid pB87T UL105. The             |
| CC | amplified HCMV UL105 orf sequence and it translational product were        |
| CC | determined. The HCMV UL105 orf was recombinated into the baculovirus AGNPV |
| CC | genome by standard co-transfection protocols. The invention relates to a   |
| CC | new human cytomegalovirus (HCMV) primase. A screening assay for compounds  |
| CC | which inhibit HCMV primase (preferably with an IC50 of not greater than    |
| CC | 200nM), comprises incubation of the compound with the primase. The         |
| CC | primase of the invention is useful as a screening tool for HCMV            |
| CC | antivirals, and as a diagnostic tool for diseases resulting from HCMV      |
| CC | infection  |
| XX |  |
| SQ | Sequence 2871 BP; 543 A; 971 C; 795 G; 562 T; 0 U; 0 Other;                |
|    |  |
|    | Query Match 1.78; Score 153.4; DB 2; Length 2871;                          |
|    | Best Local Similarity 50.3%; Pred. No. 8.5e-34;                            |
|    | Matches 376; Conservative 0; Mismatches 371; Indels 0; Gaps 0;             |
|    |  |
| QY | 585 CAATATTATTGTAATCGACGAGCGCGGAACCCCTGTCTCTTACATACTCACACCGTCGT 644        |
| Db | 765 CAATATCATCTCATCGAGTGCGCGCTTATGCTGCCTACATGCTGCAGGTGGTGGT 824            |
| QY | 645 GTTCTTTTACTGGTTCCTAAACAGCTGCTAATAATCCCCTCTTTACGCCAGGGGCGGT 704         |
| Db | 825 GTTTTTTTTACTCTTTTAAACGCCCTTGGGCGACGCGACTTTTACCAGCAACCCCGT 884          |
| QY | 705 TCCTGTCATAGTAGCGTTCGGTTTCGCCAACGCGACAACACGCTTCCAGCTCAACTTACAA 764      |
| Db | 885 GCCTTGTCATCTGCTGCTGGTTTCGCCACGCGACACCGAGCGCTGGAGAGCCCTACGA 944         |
| QY | 765 CCNGGGAGCNCNAAAGACGAGATATCGTCGTCGAGAACATCTTAATTCATGATAGG 824           |
| Db | 945 CCCTACACGCAAACAAAGAGCGTGGCNAAGGGCGTTGACGTGCTCTCGGGCTGATCA 1004         |
| QY | 825 GAAGAAGCTCGTATCTGAGTAGCTACATCTTGAGAGAAAATGGCGCTGTTTATAACAA 884         |
| Db | 1005 GAAAGAGTGCTCATCAACTACTGCGACATCCCGACAACTGGGTCATGTTTATTACAA 1064        |
| QY | 885 CAAGCGTGCACCGATCTCCAGTTGGACACCTTCTAAABATTTTACAGATATATCTTC 944          |
| Db | 1065 CAAGGTTGACCGACACTGGACTTTTGGACCTGCTCAAGTACATGAGTTCGATCCCC 1124         |
| QY | 945 CATTCCTGACAAAGTCATGAGTTAGCTAGACAGATTTCTGCTTCTCTAAAAGTAAATAT 1004       |
| Db | 1125 GCTCAAGGAGGACGCTGGCGCTAGCTGATCGCTTCGTGGCGGCCCGCCAGCTCCATCCG 1184      |
| QY | 1005 GGATCTTTTAGATATACATTGGCTGGACCCGACTCTTTTGTGCATACAGCAGGTAAGGC 1064      |
| Db | 1185 CAAACCCCTCGTACGCCGCCAGATGACGGGCTTTTCTCTTCTCAACGTCGAGGTGCAGGC 1244     |
| QY | 1065 GTATCTTAACAAACCTACACATGTCCTAACGCTAGGGGGCGATACACAGGAGNACGAGCT 1124     |
| Db | 1245 TTACTTTCAAGCGGCTCAGACGACATCCGCTGAGCGAGCGCCACCGTCTCTTTGATCT 1304       |
| QY | 1125 CTTTACCTGTCCGCTGTGGCGAGTGTTTGTGAAGCGCTTTTGGAGAAATACAGCGGGC 1184       |







CC disability or death; or asymptomatic adult infections such as  
CC mononucleosis, hepatitis, pneumonitis and retinitis. The vaccine can also  
CC be used for immuno-compromised patients suffering from acquired immune  
CC deficiency syndrome (AIDS), or undergoing chemotherapy or tissue  
CC rejection therapy after organ transplantation. The vector is also useful  
CC as a tool for studying the replication of CMV using cell-line models and  
CC the function of essential cis acting genes or nuclear acid sequences. The  
CC YAC vector provides an ease and rapidity in introducing single or  
CC multiple mutations into a CMV DNA sequence. The capacity of the vectors  
CC to generate multiple mutations exceeds the capacity of other vectors,  
CC e.g. bacterial artificial chromosome. The present sequence represents  
CC human cytomegalovirus strain AD169 genomic sequence (also see Genbank  
CC X17403.1), given in the present invention  
XX  
XX  
SQ Sequence 229354 BP; 49475 A; 64911 C; 66192 G; 48776 T; 0 U; 0 Other;  
Query Match 1.7%; Score 153.4; DB 6; Length 229354;  
Best Local Similarity 50.3%; Pred. No. 2.1e-32;  
Matches 376; Conservative 0; Mismatches 371; Indels 0; Gaps 0;  
QY 585 CAATATATTGTAATCGACGAGCGCCGGAACCCCTGCTCTTACATACCTACACCGTCGT 644  
Db 152690 CAATATCATGTCATCGACGAGTGGCGCTTATGCTGCGCTACATGCTCAGGTGGTGT 152749  
QY 645 GTTCTTTTACTGTTCTCTAAACAGCTGGCTAAATACCCCTCTTACCGCCAGCGGCGGT 704  
Db 152750 GTTCTTTTACTACTTTTACAGCCCTGGCGACACGCGACTTACCGGAAACGCGGT 152809  
QY 705 TCGGTGATAGTATGGTGGCGAAGCGGACGAGCAAAAGCGTTCAGTCAACTTACAA 764  
Db 152810 GCCTCGCATCATCTGGTGGTTCGCGCCACGACGACCGAGCGCTGGAGACGCGTACGA 152869  
QY 765 CCACGGGACCAAAAGACGGAGATATCGTGTGCGAGAACATCTTAAACATTATGATAGG 824  
Db 152870 CCATACACGCAAAAAGACGCTGGCGAAGCGGTGCTCGCGCGTATTTCA 152929  
QY 825 GAAGAAGGTGCTATCTGAGTACGTACACTTGGAGAGAACTGGCGCTCTTTTAAACAA 884  
Db 152930 GAACGAGGTGCTCATCAACTACTCGACATCGCGCAACTGGGTGATGTTTATTCACAA 152989  
QY 885 CRAGCGTGCACGATCTGAGTTCGGACACCTTCTTAAAAATTTAGAGTATTAATCTCC 944  
Db 152990 CAAGCGTTGCACGACCTGGACTTTGGCGACCTGCTCAAGTACATGGAGTTCGTTATCC 153049  
QY 945 CATTCCTGACGAAGTCATGAGTTACGTAGACAGATTGTGCTTCTTAAAGTAAGATTAT 1004  
Db 153050 GCTCAGGAGGACGACGTGGCTTACGTGATCGTTCGTGCGCGCCCGCTCCATCG 153109  
QY 1005 GATCCCTTTAGATAATATTGGCTGGACCGGCTCTTTTGTTCATAGGAGTAAAGC 1064  
Db 153110 CAACCCCTCGTACGCGCCGAGATGACGCGGCTTTTCTCTCACAGTCAAGTTCGAGC 153169  
QY 1065 GTATCTAAACAACTACACACATGTCTAAGCTAGGGGGGATACACAGGACACGAGCT 1124  
Db 153170 TTACTTCAGGGGTGACAGAGAGATCCGCTGAGCGAGCGCACCGTCTCTTGATCT 153229  
QY 1125 CTTTACCTGTCCTGGTGTGCGAGTGTGTTGTAAGCCGCTTTGAGGAATACAAAGCGGCG 1184  
Db 153230 GCCCGTCTACTGGTGTGTAACAAACCGCGGTACCAAGAGCTCTGGGAGTGGCCGACCC 153289  
QY 1185 CGTCAACTCACCACCTCACTCCGTGACGAGTGGGTGACAAAACCTCTTTAAGTTAAG 1244  
Db 153290 GCTGGGCGACTCGCCGCGAGCCGCTGAGCTCTGGTTCGCGCAGAACTTGGCGCGCATCAT 153349  
QY 1245 TAATTATTCGAGTTTGTGGACGAGGACATGTCCATAGTGCACCGAATCCACTGAACG 1304  
Db 153350 TAACCTACTCGAGTTTGTTCGACCAACACCTCTCCAGGAGATCACCAGAGGCGCTGGG 153409  
QY 1305 TTGACCCAGGTACCTTTATACCAA 1331  
Db 153410 CCCCGCGCGGACGCTGTCGTCACCAA 153436

## RESULT 15

AAH74201  
ID AAH74201 standard; DNA; 124884 BP.  
XX  
XX AC AAH74201;  
XX  
DT 15-OCT-2001 (first entry)  
XX  
XX Nucleotide sequence of the Varicella virus Dumas strain.  
XX Dumas strain; pox vaccine; vaccine; poxvirus disease; chickenpox; ss.  
XX Varicella virus.  
XX  
XX WO200156600-A1.  
XX  
XX PD 09-AUG-2001.  
XX  
XX PF 31-JAN-2001; 2001WO-JP000678.  
XX  
XX PR 31-JAN-2000; 2000JP-00062734.  
XX  
XX PA (OSAU) UNIV OSAKA.  
XX  
XX PI Gomi Y, Sunamachi H, Takahashi M, Yamanishi K;  
XX WPI; 2001-488845/53.  
XX  
XX PT Method for checking absence of mutation at specific positions of  
PT varicella virus genome for quality control of attenuated live varicella  
PT vaccine.  
XX  
XX PS Claim 1; Page 61-105; 158pp; Japanese.  
XX  
XX The present sequence represents the nucleotide sequence of Varicella  
CC virus, Dumas strain. The specification describes a method for the quality  
CC control of an attenuated pox vaccine, characterised in that the DNA  
CC sequence of specific parts of the viral genome in a sample is determined  
CC and proved to be conserved rather than mutated in comparison with the  
CC same parts of a reference viral genome. The method is useful for quality  
CC control of vaccines for use in the prevention of poxvirus diseases such  
CC as chickenpox  
XX  
SQ Sequence 124884 BP; 33789 A; 29295 C; 28177 G; 33623 T; 0 U; 0 Other;  
Query Match 1.7%; Score 152.4; DB 5; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 2.7e-32;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;  
QY 585 CAATATATTGTAATCGACGAGCGCGGAACCCCTGCTCTTACATCTCACACCGTCGT 644  
Db 96706 CAACGTTATTGTTATGATGAAGCAGGATTCCTAGGCGGTCATTTCTCAGCGCGTGT 96765  
QY 645 GTTCTTTTACTGTTCTCTTAAACAGCTGGCTAAATACCCCTCTTTACCGCCAGGGGCGGT 704  
Db 96766 TTACTGTTGGTGGCTTTTGAATGCTATATATCAAGGCCCTCAGTACATAAACGGTCA 96825  
QY 705 TCCGTGCAATAGTATGCTGCGGTTGCCACGACGACAAACGCTTCCAGTCAACTTACAA 764  
Db 96826 ACCGCTCATAGTATGCTGCGGTTGCCCAACCTGACTCGTTAGAACTCTCATTTTCA 96885  
QY 765 CCACGGGACGAAAGACGAGATATCGTGTGGGAGAACATCTCTAAACATTCTAGTATAGG 824  
Db 96886 ACATGACATGACGGTTTACACGTTACTCTTAGTGAAATATCTACGATATAATCTG 96945  
QY 825 GAAGAAGTCTGATCTGAGTACGTTTGGAGAGAACTGGGCGCTGTTTATAACAA 884  
Db 96946 CAATCAAACTCTGGCTCAATATCTAAACATCTCAATCACTGAGTGGGCAATCTTTTATTAATAA 97005  
QY 885 CAAGCGTGCACCGATCTGCAAGTTCGGACACCTTCTAAAAATTTTAGAGTATTAATCTTCC 944  
Db 97006 CAACGATGTCAGAGGACGATTTTGGAAATCTTTTAAAGCGTTGAGTACGGGCTACC 97065

|    |       |  |       |
|----|-------|--|-------|
| Qy | 945   | CATTCTGACGAGTCATGAGTTACGTAGACAGATTGTGCTTCCTAAAGTAAGATTAT | 1004  |
| Db | 97066 | TATTACCGAAGCACATGCGCGTCTGTCGATACATTGTGTACCTGCTCCTATATTA  | 97125 |
| Qy | 1005  | GGATCCTTTAGAATACATTGGCTGGACCGACTCTTTTGTACATAGCGAGGTAAGGC | 1064  |
| Db | 97126 | CAATCCTGCTAATCTTCCGGATGGACCGCTCTGTATTCTGCATAAGGAGGTGAGCG | 97185 |
| Qy | 1065  | GTATCTAACAAACCTACACACATGTCTAACGCTA                       | 1098  |
| Db | 97186 | GTATATGAGTAAGTTACACGGCATTTTAAACTA                        | 97219 |

Search completed: April 10, 2004, 22:00:17  
Job time : 2133 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 20:55:09 ; Search time 393 Seconds  
(without alignments)  
12811.882 Million cell updates/sec

Title: US-09-831-000-1\_COPY\_60966\_70038  
Perfect score: 9073  
Sequence: 1 atggagagtcctcgatg.....ggggccacgagtcgcttga 9073

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq\*  
3: /cgn2\_6/ptodata/2/ina/5A COMB.seq\*  
4: /cgn2\_6/ptodata/2/ina/5B COMB.seq\*  
5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq\*  
6: /cgn2\_6/ptodata/2/ina/backfileseq1.seq\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length  | ID | Description          |
|------------|-------|-------------|---------|----|----------------------|
| 1          | 1002  | 11.0        | 35100   | 2  | US-08-770-379-18     |
| 2          | 1002  | 11.0        | 35100   | 3  | US-08-757-669A-18    |
| 3          | 1002  | 11.0        | 35100   | 4  | US-09-230-371A-18    |
| 4          | 714   | 7.9         | 35100   | 2  | US-08-770-379-19     |
| 5          | 714   | 7.9         | 35100   | 3  | US-08-757-669A-19    |
| 6          | 714   | 7.9         | 35100   | 4  | US-09-230-371A-19    |
| 7          | 154   | 1.7         | 125157  | 4  | US-09-913-514-2      |
| 8          | 152.4 | 1.7         | 124884  | 4  | US-09-661-596A-76    |
| 9          | 152.4 | 1.7         | 124884  | 4  | US-09-913-514-1      |
| C 10       | 137.6 | 1.5         | 1532    | 4  | US-09-651-656-36     |
| C 11       | 137.6 | 1.5         | 1532    | 4  | US-09-650-855-36     |
| C 12       | 134   | 1.5         | 723     | 4  | US-09-489-039A-4593  |
| C 13       | 125.8 | 1.4         | 7316    | 3  | US-09-289-254-3      |
| C 14       | 117.6 | 1.3         | 1830121 | 4  | US-09-557-884-1      |
| C 15       | 117.6 | 1.3         | 1830121 | 4  | US-09-843-990A-1     |
| C 16       | 111   | 1.2         | 4161    | 3  | US-08-790-517-1      |
| C 17       | 109.4 | 1.2         | 4161    | 3  | US-08-750-517-1      |
| C 18       | 94    | 1.0         | 738     | 4  | US-09-252-991A-13238 |
| C 19       | 91    | 1.0         | 119211  | 4  | US-09-596-002-40     |
| C 20       | 90.2  | 1.0         | 657     | 4  | US-09-134-001C-2483  |
| C 21       | 87.8  | 1.0         | 768     | 4  | US-09-543-681A-3433  |
| C 22       | 84.2  | 0.9         | 640681  | 4  | US-08-790-988-1      |
| C 23       | 84    | 0.9         | 8753    | 4  | US-08-426-630-2      |
| C 24       | 80.6  | 0.9         | 564     | 4  | US-09-107-532A-353   |
| C 25       | 74.8  | 0.8         | 729     | 4  | US-09-328-352-3484   |
| C 26       | 63.4  | 0.7         | 3662    | 4  | US-09-221-017B-808   |
| C 27       | 61    | 0.7         | 17846   | 4  | US-08-956-171E-109   |

|      |      |     |         |   |                     |
|------|------|-----|---------|---|---------------------|
| C 28 | 50.6 | 0.7 | 504     | 2 | US-08-747-887-22    |
| C 29 | 59.2 | 0.7 | 4403765 | 3 | US-09-103-840A-2    |
| C 30 | 59.2 | 0.7 | 4411529 | 3 | US-09-103-840A-1    |
| C 31 | 49.4 | 0.5 | 19124   | 2 | US-08-487-826B-13   |
| C 32 | 49.2 | 0.5 | 11049   | 4 | US-10-204-708-23    |
| C 33 | 46.6 | 0.5 | 5152    | 4 | US-10-204-708-74    |
| C 34 | 45.6 | 0.5 | 505     | 4 | US-09-621-976-15639 |
| C 35 | 44.2 | 0.5 | 6801    | 4 | US-10-204-708-61    |
| C 36 | 44.2 | 0.5 | 6865    | 4 | US-10-204-708-20    |
| C 37 | 42.8 | 0.5 | 5852    | 1 | US-07-867-106-2     |
| C 38 | 42.8 | 0.5 | 9347    | 4 | US-10-204-708-35    |
| C 39 | 42.4 | 0.5 | 1230025 | 4 | US-09-198-452A-1    |
| C 40 | 42   | 0.5 | 6113    | 4 | US-10-204-708-14    |
| C 41 | 41.8 | 0.5 | 474     | 4 | US-09-621-976-18033 |
| C 42 | 41.8 | 0.5 | 53332   | 4 | US-09-801-861-3     |
| C 43 | 41.8 | 0.5 | 640681  | 4 | US-09-790-988-1     |
| C 44 | 41.6 | 0.5 | 1001    | 4 | US-09-671-317-141   |
| C 45 | 41.2 | 0.5 | 3740    | 4 | US-09-489-847-14    |

ALIGNMENTS

RESULT 1  
US-08-770-379-18  
; Sequence 18, Application US/08770379  
; Patent No. 5849564  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Echenzky, Roy A.  
; APPLICANT: Russo, James J.  
; APPLICANT: Edelman, Isidore S.  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED  
; VIRUS: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/770,379  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 52342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35100 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-770-379-18

Query Match 11.0%; Score 1002; DB 2; Length 35100;  
Best Local Similarity 54.5%; Pred. No. 5.1e-292;  
Matches 2414; Conservative 0; Mismatches 1765; Indels 253; Gaps 10;  
QY 1 ATGGAGAGTTCGTCGGATGGACCAACACGTCGACCAATCCGGGGTTCATTCTTGAAC 60

|    |       |  |       |
|----|-------|--|-------|
| Db | 29792 | ATGGACAGCTCGGAAGGTCGCACTGATATGGACGAGCCAGCCCGCGGTTCATCTCAAC     | 29851 |
| Qy | 61    | ATGACGTCGATCCCAAGTCAGGGGTGTCGNGATACGTCAGTCGCTGTCGAATATA        | 120   |
| Db | 29852 | ATGACATCGATGCCAAGTCAGGAGCGTAGTGGACAGATTGACAGGTTGTCAATATC       | 29911 |
| Qy | 121   | ACTACAGCCCAACCGGAATGGTTTGGTACGACCTGGCTTCGATCCGGCTGAAGACTCC     | 180   |
| Db | 29912 | ACTACCTCGCCCGGAGATGGCTGGTATGACCTAGAGTTGATCCACTGGAAGACAA        | 29971 |
| Qy | 181   | GGSCCGTCTTCGCGTTTACCTTTATCTAAATTAGGGAACCTGCTGGTGGGAAAGT        | 240   |
| Db | 29972 | GGCCCTCTTCGCGTTTTCGGCATACGTAATACGGGAGCTGACAGGAGCGGGAAAGC       | 30031 |
| Qy | 241   | ACCAGCATATCGGCCCTGTACCAAAATTTAAACTGCCCTGATCAAGGGCGCACCACTA     | 300   |
| Db | 30032 | ACCAGGATATCGCCCTACATCAGAACTCAACTGCCCTAAATTACGGGGCTACAGTGTA     | 30091 |
| Qy | 301   | GGCGCACAACTTACGCGTCGCTTAAGAGTTCTGTCGCAAGATCTTCAGCGCTTTT        | 360   |
| Db | 30092 | GGCGCACAACTTTTCAGGGCTTTAAAGTCTTACTGTGCCATATATACCGCCTTC         | 30151 |
| Qy | 361   | GGTTTAAAGCGGACACATCAATATAGCCGTCAGAAAGCTCATACAGCGGACCGCTA       | 420   |
| Db | 30152 | GGATTCAAGAGCAGACACATTAATATCTGCCAGAGGAAAGTGCCTCAAGTAACTCAGTCC   | 30211 |
| Qy | 421   | TCCATAGACAAATTCAGCAACAGAGCTATCGAAGTATTCGCGGCTTAGTGACATT        | 480   |
| Db | 30212 | TCCATCGCAACTCCAGAGATACGAGCTGGCTAGTACTGCGCACTGTCACCCATTT        | 30271 |
| Qy | 481   | ATGAAGAGGTTATGGCGAAATAAACCCTATGCGATGTACGGGACTATCAACCGCAAT      | 540   |
| Db | 30272 | ATTCGAGAAATTTATGCGAAGAAACAAAGGGGCGAGTATAGCTCCCTCTCTCAAGCGCT    | 30331 |
| Qy | 541   | TTTGAACCCCTCTCGAAGATGACCGGACCGTGTATATGAGCTTCCATATTTATGTAATC    | 600   |
| Db | 30332 | TTCAGACTCTTTGCGGTATGGTGGAGCAATTTGTCGACGAGTACATTTATCGTGATA      | 30391 |
| Qy | 601   | GACGAGCGGAAACCTGTCCTTACATATCATACACCGTCGCTGTTTACTGTTTC          | 660   |
| Db | 30392 | GACGAGCTGGAACCTCTCTGCCATATTTTGAAGCGGCTGCTTCTTATGTTT            | 30451 |
| Qy | 661   | CTAAACAGCTGGCTAAATACCCCTCTTTACGCCGACGGGGGGTTCGCTGCATATATGC     | 720   |
| Db | 30452 | TACACAGTTGCTGGACACCCCGCTATACAGAAATGTCGCGCTGCTTGCATATGCTGC      | 30511 |
| Qy | 721   | GTGCTTCGCCAACGACAGAAACCGCTTCCAGTCACTTACACACGAGGACGCAAAAG       | 780   |
| Db | 30512 | GTGGGCTCTCCACCCAGAGGAGCGCTTTCAGTCGGTCTTCAACCAACGACAGAGA        | 30571 |
| Qy | 781   | ACGGAGATATCGTCGTGGAGAACATCTTAATTCATGATAGGAGAGAGTCTGATCT        | 840   |
| Db | 30572 | AACGAGATATCGCTGTGATATGCTGCTCACTTCTTATGGGAAACGTCAGGTTGCA        | 30631 |
| Qy | 841   | GAGTACGTACACTTGGAGAGAACTGGCGCTGTTTATTAACAAACAGCGCTGCACCGAT     | 900   |
| Db | 30632 | GATTATATAGGCTGGAGAGAAATGGGCGCTTATTAACAAATAAGCGCTGTACGGAT       | 30691 |
| Qy | 901   | CTGAGTTGCGACACTCTTAATAATTTAGATATATCTTCCATTCCTGACGAGTC          | 960   |
| Db | 30692 | CCCCAGTTGGTCATCTGCTGAAGACCTTAGAATATATCTAGACATATCACAGAGTTA      | 30751 |
| Qy | 961   | ATGAGTTAGTACAGATTTGTGCTTCTTAAAGTAAGATTTAGATCTCTTTAGAAATAC      | 1020  |
| Db | 30752 | ATGGACTATATAGATAGTTTGTGGTTCCGAAGAGTAAGATTTCTGGACCCGCTCAGTAT    | 30811 |
| Qy | 1021  | ATTGCTGACCCGACTTTTGTTCATAGCAGGCTAAGGCGTATCTAAACACCTTA          | 1080  |
| Db | 30812 | GCAGGTTGACAAAGACTCTTATCTCACACGAGGAGTGAAGTCTTTTCGGCAACGCTG      | 30871 |
| Qy | 1081  | CACACATGCTAAACGCTAGGGGGGATACCAAGGACAGAGCTCTTTACTGTCGCTG        | 1140  |
| Db |       |  |       |
| Db | 30872 | CACACCTGCTGTCGAGTAATAAGATGCTGTCTCCACAAAGCTTTTACCTGCCAGTG       | 30931 |
| Qy | 1141  | GTGTCGAGGTGTTTGAAGCCGTTTGAAGAAATACAAAGCGGCGCTCAACCTCACCAAC     | 1200  |
| Db | 30932 | GTCTGTGAGGTGTTTACAGCCATTTAGAGAGTACAAACGGGCGTAGGCTCACACAC       | 30991 |
| Qy | 1201  | CTCACCGTACCGAGTGGGTGACAAAAACCTCTTTAAGTTAAGTAATTTTCCAGTTT       | 1260  |
| Db | 30992 | ATGACTCCCATAGAAATGGTTAAACAAAATCTTTTACAGGCTAAGTAATCTACTGCACTT   | 31051 |
| Qy | 1261  | GTGGACACGAGATGTCATAGTCCCAAGATTCACCTGAACTGAACTCCGACCCAGTTACC    | 1320  |
| Db | 31052 | GCTGATCAGGACATGGCTGTGTTGGACCTTATATCACAGACGCTCCACAGATCACC       | 31111 |
| Qy | 1321  | TTTATCACCAAGTTTGTPTAAAAACAGCACGCTGCTTTTAAATGAAAAACAAAAATGC     | 1380  |
| Db | 31112 | TTTCGCCATAAATTTGTCAAAAACAGCTATGCTACCTTACTTGAAAGACCAAAAAATGT    | 31171 |
| Qy | 1381  | ATATGTCGGTTTACGGTACGTAATTTTGTAGTTTCAAAAGATCTTAGACAGGACTCTTC    | 1440  |
| Db | 31172 | ATATCGGGTTTACGGGTATACCAAGATTTCAAGTCCATCTTAGACGGGAGCTATTT       | 31231 |
| Qy | 1441  | GTGGAGACGCAATTCGACAGACGCTCCGGAATACGTATATACGTTTCTTAAACACGCTACTG | 1500  |
| Db | 31232 | ATCGAAAGTCAATTCGACATAACCCGCTTATGTGTACAGTTTCTTAGTACCTGCTA       | 31291 |
| Qy | 1501  | TACACGCCATGTACTCGTTTACGCGTACGGCTGACTAGGTACACACGAGAAATCTG       | 1560  |
| Db | 31292 | TATATGCCATGTACTCATTTTACGCAAGGGTGAAGCAGGGGCATGAAGATTCCTC        | 31351 |
| Qy | 1561  | CAAGACCTTAAATTTGCAACCCCTCCCGCGCTCTGGCAACCGGGCGGTGAGACCTTCAA    | 1620  |
| Db | 31352 | AGGGACCTTCAGGGAATGCGCGTCTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGAC    | 31411 |
| Qy | 1621  | ACGGTTCTGTAAGAGTTTAAACCTCGAAGACGACATCTTACACGCTGCTCAGTCCCGCG    | 1680  |
| Db | 31412 | GTTCTGGGGCAGGAG-----GGGGACACAGATGCCCTTCTACCTCACCGCAGCTCCCA     | 31465 |
| Qy | 1681  | CAACCGCGGTATCACCTCCCTCCAGGTTTGTGTGACACAGTATGTCGCGCTTAAAGGAC    | 1740  |
| Db | 31466 | CCATCCCGCACCCACGGGCTCTTCCAAACATGCTGGTGCCTTACTCCGGGGCAAGAA      | 31525 |
| Qy | 1741  | GTGTTGCGCTCCAGAAATAAGGTGCGTGTGCTGTTTCGGGGGAGTTTGAGAGGAA        | 1800  |
| Db | 31526 | CTATTTGCAACAGGTGGCCCTGGACGCCGACATTTTGTGACGAGTTCTTCCACTCC       | 31585 |
| Qy | 1801  | ACGTTTTCGCGTTTACCGTTTAAATGCTGTAAGGACGAGTTGACTTTGTCTCCCT        | 1860  |
| Db | 31586 | GATTTTCAACGTTTACCGTGAAACATCGTGTGCGAGATGGCGTGGACTTTGTGTCCACT    | 31645 |
| Qy | 1861  | TCAGAACGCTTCAACGGGCTGTTGGCTTTGCTATCGACCGTTCGATCGTATTAATTAAG    | 1920  |
| Db | 31646 | TCCCGGGCTCCACGGCTAGTGGCATACGATCCACTATAGACCTATATATATCCAG        | 31705 |
| Qy | 1921  | GGGTACACGTTTTCACCGGTAGCGTTCGCTCGCTGTGAGGGTTTGGCACTCAGTGATGAC   | 1980  |
| Db | 31706 | GGATATACGTTTCTCCCAAGTGAATTCGGCGCTCCAGGAGGACAGCGCTCAGCGAGAC     | 31765 |
| Qy | 1981  | CTCAGGAGAGATGCTCCCTCCCTGCTGTCAGGACTCTAGCGGTTTATCGGCTGCTA       | 2040  |
| Db | 31766 | CTGGCAGAAAGATGCCCTCATAGTTGTCAGGACTCATCGGGTTTCAITGCTGCTG        | 31825 |
| Qy | 2041  | GAGAAATACATAACCAAAATTTGACGAAACCATGAGGACGGGAGCGTTTCCAAAGTGTGC   | 2100  |
| Db | 31826 | GAAATAACGTCACCAAGATGACAGAGACCTCCGAAGGTGGCGACGCTGTTTAAACATATGT  | 31885 |
| Qy | 2101  | TGTCGGGGGACTATGGGTGAGTCAAAATTTTAGCCATGACCATCTGTAAGGACAGGGA     | 2160  |
| Db | 31886 | TGTCAGGGGACTACGGTATCAGTTCTATCTGGCTATGACCAATAGTAGGACAGGGG       | 31945 |
| Qy | 2161  | ATGCTGTGGAGGAGTACCGGTAGTATTTGGGTCCCAAGAACGTCACAGAACCCAC        | 3220  |
| Db | 31946 | GTTTCACTAAGTAGGGTGGCCATATCGTTCCGCAACCAACCGCAATATACAGAGCAGTCTA  | 32005 |

QY 2221 GTGATGTAGCAATATCAAGGCTGTAACTCAAAATTTTGGTTCATGACAGCAACCCC 2280  
Db 32006 GTGATGTGGGTGTATCCAGGSCCATCGACGCTCGTTACCTGGTAAATGGAAGTAATCCC 32065  
QY 2281 CTTAAACCCCTCCTCAGAGAACCACTCGATAAACAACCTCGGCAAGCATATAGTCGGGCC 2340  
Db 32066 CTTAAAGCTAATGGACCGCGGTGACGCCCAAGTCCCATCCTCAAAGTACATCAAAAGCC 32125  
QY 2341 CTTCCACACCCCAACACACACCTCATCTACTAA----- 2373  
Db 32126 CTATGCAACCCCAAGACTACTCTGATCTACTGACCGGTACCCCTCTCTTAGGACACTGAT 32185  
QY 2374 -----ATAAATAAATAAAGCAAAATGAAACAAATTTG 2406  
Db 32186 GTGTTTGGGAATAAAGCATGAGACTTGACACCTATAATGGTCTGTATTGACACCAATCTT 32245  
QY 2407 TTTTATTTCAGTCCAAACAGGCGCAGGATGTTGTCATTTCACACACACCGGGGGTGG 2466  
Db 32246 TTATTATCAGTCCAGCCAGGCGCAATATATGCAACCGTTTCCACACAGGGGTGGCGTGG 32305  
QY 2467 CGCCCATAGTTTGACACCAAGAAACCGCCGGCCCTCTGGGTTAGAAGTAGCCCC----- 2521  
Db 32306 AGGCCAGATGCGGGTGGGTGCGTGCACCTGACACCCCGCGTGTGTGTCTCTGATG 32365  
QY 2522 -----GCCCCCGGTGTCGAGGTGATACGTCTCTTAGGGCCCTTGGTT 2564  
Db 32366 AAATCAGTGGCGGAAGTACTGGGAGATTGGGTTGGGAGGTGACCCCTTTGTCGACG 32425  
QY 2565 GTACCAACATAATACGTCGTATCACCGTTGCACTCCCTCGTGAACCGCGCCCTCTG 2624  
Db 32426 GAGACAGATCAGCTCAGCGGGAGAGGGTCTCTGCTGTGTCTACTCCCGAGGATA 32485  
QY 2625 CAGCGCGCAGAACACCGCGCTTGTAACTGATCT-----GCA 2664  
Db 32486 TAATTATCAGGACGCCACTGCTTTGCGGCTTAAGTTTGGTGTCTCTGGCAGCGACCA 32545  
QY 2665 CGTGTGCTGTCTGAAGATGAACACAGGTCAATCATCATGAGCGCGCCGCTCAT 2724  
Db 32546 CATCTCGTACAGAGGAGGGGTAGATGCGCTTTTGGGCTTTGGGCCACGTCAATGA 32605  
QY 2725 CGCGCTGTATTCGTCCCATGTCTCAACAGAGCTCGGTGCTCACTTCGTGTCATCTG 2784  
Db 32606 GCCGATCTCTGACTCAATCTTCCCTTGGTCTTCTCGTCTCTCGGACGAGGTG 32665  
QY 2785 AATAATAATATGCTGTCTGCTGAGTGTATCCAGCAACACAGGTCTTCTTCTCATCG 2844  
Db 32666 GCTGTGGGAAAAATGGCGCGCTCGGTAAACCGCGCTCATTTGTTTACGTCGGAGAT 32725  
QY 2845 AAGACCATCGCCCTCTGCGGTGATCTAGATATATCATCTGACAGGCGGCGAGGCC 2904  
Db 32726 TGGAACTGTATCGCTATCAGATCCGATGTCAAGTCGAGTCGCGGTGGGTGCGCGC 32785  
QY 2905 G----- 2905  
Db 32786 GCAGGGGGCGCACAGGGCCCTTATCAGGGTCTGTATGTTGTAACCTTTGTGTTCCAG 32845  
QY 2906 -----TGGCGAAAGCCGTCAGCTTCTCGCCACTTTTCATCAGTGT 2946  
Db 32846 GTACACTATTTCTGGAAGCAGGTGAAAGTCCGATGCCCCGGTCCAGTGTATGCCGCCA 32905  
QY 2947 CCTCGTGTATCTATAAAAAATTCGTTCTGCTGTCTGATTCAGATGACTGGGGTGGC 3006  
Db 32906 TCGGTTCCAGGATAGCAACCCCTCGTCTGTGAAGGTGAGGCCAGCAGGCGGAAATC 32965  
QY 3007 GGTGTGGCGAGTAACACCGACGACGGC-----CGCGACCCACCAAGCG 3053  
Db 32966 CGTCATCTGACTAACCCATCCCATGACCGCTCGGACTCGCGGTGTCGTTGAACTGC 33025  
QY 3054 GTCAGTGGCGGGCGCTCGGTGTGTGTGGCACCTGTGCAAGTAATGACGGTAGGT 3113  
Db 33026 CACCGCGCCGCTACCACTGCTACCGTTTGGGCGTATGGGCCGCTCTGGCCAGAGGCC 33085

QY 3114 AGCGCGCGCTCTGGTTAAAG-----ACTCACCGTCCGGTG 3150  
Db 33086 TCGGCGCAAGTGAAGTTGAAAAAGTCTGCAGGGTACCCTCTGGCTCGTCTT 33145  
QY 3151 CGTCCGATCTATTCTAGTGGGAGAGACCGCGGTTCTCTTCTTCAATCCACATCAC 3210  
Db 33146 CTTCTCGAAATCATCGTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 33205  
QY 3211 TGAGTCTGCTGTTCTCCGAATGAGTCCGTTTCTGCTGGCGTTTGGAGTAGTGGGAG 3270  
Db 33206 ATTCGCGCTCGACTGATCCGGGATATCTGTAGATCCAGAGGGTTGCTGGCGGATG 33265  
QY 3271 GGGATCTTGGTCAAAACATCTCAGGTGACGTATGTAATCAATCAGTGGGGGCGCTCCGT 3330  
Db 33266 GCGTCTCTCGGGAAGACGCTCTGCGGACAGATATCTATCAACGCTGGGTCCAGCAT 33325  
QY 3331 ACCCGCGGTATCCGGCAACCCCGTACGGG-----CCCGGCGAAACAAAAAC 3378  
Db 33326 AGCGCGCGCTGCCAAATCTCTGGAAGTATGAAGAGGTGAGGTGGGAATATGAAT 33385  
QY 3379 GTCTTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3438  
Db 33386 TCACGGGGGTCTGCTGCGAGGCGCTCTTCAATTGGAAGCATTTCTCTTCTTCTG 34445  
QY 3439 GGGTCTTGAACGGGTCCGATAAGAAC-----ATAGCCATCGCC 3476  
Db 3446 TGCTAGACGAGTCTCTCAACATCGCATGCGCTTGTACGGGTTGACCGCTAGGGC 33505  
QY 3477 GAAAGTCTACCTGTAAACGACTGCTATCCCGATCAGTTAATATTCCAGTCAACGG 3536  
Db 33506 GGAATTTACAAACACACGAGTTATTGCTCTTACTCTCCACAGGCCCGAGTCCACAG 33565  
QY 3537 CACCCCTCCGATCTGAATAGTATGTCATTTCGCAAGTTTAAATGGTTGAGCCTAGAA 3596  
Db 33566 TCTCAGCGCGTGGCGAGTCAATAGTCTTGGTCTTAAAGTATTAAGTATTAAGCCTTGA 33625  
QY 3597 ACTCGGCGAGGAGAACCGGTTCCCGTGGGCGGTAGCCAGGGGAGACGATGGCGCG 3656  
Db 33626 ACCGAGCCATCGCGAGTGTCCGCCACCAAGAGAG-----GCCAGCGGAGTGGATGCTGG 33682  
QY 3657 CTTGAGGACGAGATGCTCTGTCGCTTATACGCTGCGCTCTCTATAGCTTTCTTCT 3716  
Db 33683 CCGTAAGCACGAGTGTCTGCTGGTTATGAGCGGAGTCTCTCAATGGCTTGGCC 33742  
QY 3717 CCCACAGCATAAAAACGACCTCTGAAGCTTGGCAGAAAGCAGCTTATATGATGCTGG 3776  
Db 33743 CCCACAGGAAACGCAATGTTCTAACTTTGAGGATATGCTACTGATGATAACTCG 33802  
QY 3777 TGAACCAATCCAGCAAGTGGCGTGTGACCCCGCTTCCCTCTTCCACCGTCAGAA 3836  
Db 33803 TGAACCAATCCAGCAAGTCCCTCGTGTGAGCGCGCCCTCCCTTCTCCACCGTCAAAA 33862  
QY 3837 TGGTGTAAAGCAGAACTCCCGTTTGGCCCAACAGTCTTAAGCACCCGCTGAGAGGAG 3896  
Db 33863 CTGTGTTTGTAGCAACACACCTGGCGAGCCAGCTGTGAGGACACCCGTTGGAGGAG 33922  
QY 3897 CGGTGAACCCAGGTACAGTATTTGGTATCTCTTTTAAATAATTTTGGGCTTGGAGTA 3956  
Db 33923 TACTGAAATTTGGGACGGAAGCTCTAGCTCTCTAAAGATGCTTCTCTCTCTCTCTCTCT 33982  
QY 3957 TACTGTAGTCCGACACGCTTAAAGGTAGTCCGCTGCTTGAACCGGATGCTGTAAGGAT 4016  
Db 33983 CCGTACATTTGGGATTCACACTTAAACGCGCAGGCTAGCTTGGCCCTTGTGTACGGGT 34042  
QY 4017 CTTGCCCAAGGATCAGCTTTAATATCTCTCGGTTGCGAAATATACACACCAACATATA 4076  
Db 34043 CTTGCCCTAAGATCACCACTTTAATATCTCTGATTCGAGGAGTGGGACCAACATCA 34102  
QY 4077 TCTATCTATCGGTGATTAATCAGGTACGTTTGTCTCATATCCATAACCGGCTTCAGCA 4136  
Db 34103 GCTTGTCTGTGGGATACACTGTGGTGTGTAGCTTAACTTCCGATCTGTCTGAGCA 34162  
QY 4137 GAGCGCGAGCTTCTGTTTAAAAATGGAGAAAGATTTAAAAATTCAGGCCA 4188

Db 34163 GCGAGACAGTCTTCTGTTTCAAGAAATGATGAGAGGCTCAGAAAGGAATCCA 34214

RESULT 2

US-08-757-669A-18 Application US/08757669A

Sequence 18, Patent No. 5183751

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James J.

APPLICANT: Edelman, Isidore S.

APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS

TITLE OF SEQUENCES: SEQUENCES AND USES THEREOF

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham LLP

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/757,669A

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 29,678

REFERENCE/DOCKET NUMBER: 45185-F

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 18:

SEQUENCE CHARACTERISTICS:

LENGTH: 35100 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-08-757-669A-18

Query Match 11.0%; Score 1002; DB 3; Length 35100;

Best Local Similarity 54.5%; Pred. No. 5.le-292;

Mismatches 2414; Conservative 0; Indels 253; Gaps 10;

Matches 2414; Conservative 0; Indels 253; Gaps 10;

QY 1 ATGAGAGTTCGTCGGATGACCAACACAGTCGAAACCAATCCGGGGTTTCATCTTGAAC 60

Db 29792 ATGACAGCTCGAAGGGTCACTGATATGACGAGCCAGCCCGGGTTTCATCTCAAC 29851

QY 61 ATGACGTCGAGTCCAAAGTCAGGGGTGTCGTGGATCAGTCAGTCGCTGTCAAATATA 120

Db 29852 ATGACATCGATGTCGCAAGTCAGGAGGTAGTGAACAGATTGACAGGTTGTCAAATATC 29911

QY 121 ACTACGAGCCACCGAAATGGTGTGTGACGCTGCGCTTCGATCCGGCTGAAGACTCC 180

Db 29912 ACTACGCGCCCGAGATGGGCTGTGTATGACCTAGATTGATCCATCGAAGACGAA 29971

QY 181 GGGCGGTTCTTCCGTTTACCGTTTATCTAATACGGGAACCTGCTGTGTCGGGAAAGT 240

Db 29972 GGGCGGTTCTTCCGTTTTCGGCATACTAATACGGGAGCTGACGGAGCGGGGAAAGC 30031

QY 241 ACCGACATATCGCCCTGTACCAAAATTTAACTGCTGATACAGGCGCGGACCAACATA 300

Db 30032 ACCAGCGTATCCGCCCTACATCAGAACTCTCACTGCTAATACGGGGGTACAGTGGTA 30091

QY 301 GCGCACAGAACCTATCGCGTCGCCCTAAAGACGTTCTGTCCCGAGATCTTCAGCGCTTTT 360

Db 30092 GCGCACAGAACTCTTCCAGGGCTTTAAAGTCTTACTGTGCCACTATATACCGCGCTTC 30151

QY 361 GGCCTTTAAGAGCCGACACATCAATATAGCCGTGAGAAAGCTCATACAGCGGAGCGTA 420

Db 30152 GGATTCAGAGACGACACATTAATATCTGCCAGAGGAAAGTGCCCAAGTAATCACTAGTCC 30211

QY 421 TCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTGGCGGTTATAGTGGAAT 480

Db 30212 TCCATCGAGCAACTCCAGAGATAGAGCTGGTGTAGTACTGGCCAACTGTCCCGATATT 30271

QY 481 ATGAAGAGGTTATGGCGGAAAAAACCCTATGGCATGTACGGGACTATATCAACGCGAAT 540

Db 30272 ATTCAGAAATTTATGCGCAAGAAACAAAGGGGAGTATAGTCCCTCTCTCAAGGGCT 30331

QY 541 TTTGAAACCTCTCGAAGATGACCGGACCGTGTATATGGACTTCCATATATATTGTAATC 600

Db 30332 TTTGAACTCTCTTGGCGTATGGGTGGAGCAATTTGAGGACGAGTAACATTAATGTAATA 30391

QY 601 GACGAGCGGGAACCTCTGCTCTTACATATCTACCAACCGTCGTCGTTCTTTACTGTTTC 660

Db 30392 GACGAGCTGGAACCTCTCGTCCCATATTTGAGCGCGCTGGTGTCTTCTATTGTTT 30451

QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTACCGCAGGGGGGGTTCGTCGATAGTATGC 720

Db 30452 TACAACAGTTGGCTGGACACCCCGCTATACAGAAATGGTGGCGCTTGCATAGTCTGC 30511

QY 721 GTCGTTCCGCAACGACAGCAACCGCTTCAGTCAACTTACCAACCGGACGACCAAAAG 780

Db 30512 GTGGGCTCTCCACCGACGAGCGCTTTCAGTGGCTTCAACCAACGACGACGAGAGA 30571

QY 781 ACGGAGATATCGTCGTGCGAGAACATCTTAACATTCATGATAGGAGAGAGGTGCTATCT 840

Db 30572 AACGAGATATCTGGCTGTGATTAATGTCTCACCTTCTTATGGGAAACGTTGAGTTGCA 30631

QY 841 GAGTACGTACACTTGGAGAGAAATGGGGCGCTGTTTATAAACAAGCGCTGCACCGAT 900

Db 30632 GATTATATAGCTGGACGAGAAATGGGCCCTATTTTAAACATTAAGCGCTGTACCGAT 30691

QY 901 CTGCAAGTTTCGACACCTTCTAAAAATTTAGAGTATATCTTCCATTTCTTGACGAGATC 960

Db 30692 CCCAGTTTGGTCACTTGTGAGACCTTAGAATATAATCTAGACATATCACGAGAGTA 30751

QY 961 ATGAGTTAGCTAGACAGATTTGTCTTCTTAAAGTAAAGTATGATGATGCTTTTAAATAC 1020

Db 30752 ATGGAATATATAGATAGGTTTGTGTTCCGAGAGTAAAGTCTTGACCCGCTCGAGTAT 30811

QY 1021 ATTGCTGGACCCGACTCTTTTGTTCATAGCAGAGTAAAGGCGTATCTAACAAACCTA 1080

Db 30812 GCAGGGTGCACAGACTCTTCTCATCTCAACAGGAGGTGAAGTCTTTTCTGGCAACGCTG 30871

QY 1081 CACACATGTCTAACGCTAGGGGGGATACACAGGACACAGAGCTTTTACCTGTCCCGTG 1140

Db 30872 CACACCTGCTCGAGTAATAAGATGCTGTGTCACAAAGCTTTTCCCTGCCCAGTG 30931

QY 1141 GTGTCGAGGTGTTTGTGAAGCGTTTGAGGAATACAAAGCGGGCGCTCAACCTCACCAAC 1200

Db 30932 GTCTGTAGGTGTTTACAGGCCATTTGAGGAGTACAAACGGCGGTAGGCTCACACAC 30991

QY 1201 CTCACGTCGAGCGGTGACAAAAACCTCTTTAAGTTAAGTAAATATTTCGAGTTT 1260

Db 30992 ATGACTCCCATAGATGGGTAAACAAAAATCTTTTACGGCTAAGTAATCTCTCGAGTTT 31051

QY 1261 GTGACACGAGCATGCTCATAGTCGCCACGGAATCCACTGAACGTTCCAGCCAGGTTACC 1320

Db 31052 GCTGATCAGGACATGCTGTGTGTTGGGACCTATATCAAGAGCGGTCCACACAGATCACC 31111

QY 1321 TTTATCAACAAAGTTGTTTAAAAACAGCCAGCGTGTCTTTAATGGAAGAAACAAAAATGC 1380

Db 31112 TTTCCGCACTAAATTTGTCAAAAACAGCTATGCTACCTTCTTGGAAAGACCAAAAAATGT 31171

QY 1381 ATATGTGGTTTCAGGGTACGTACTTTTGTGATTCAAAAGAAATCCTTAGACAGCGAACTCTTC 1440





3379 GTCTTGTTCTCTTTTCCTAGTGGCCCGGAATCGGAGCATCTCTGGAGTGGCCGCG 3438  
3386 TCACGGGGGTGCTCTGCGAGCGCTCTTCAATTTGGAGGATCTCTCTTCATCGTGTG 3445  
3439 GGTCTCTTGAAGCGTCCGGATAGAAC-----ATAGCATGGCC 3476  
3446 TGTAGACAGGCTCTCAAAACATCGCCATGGCCCTTGTACGGGGTTGACCGCTAGGGC 3505  
3477 GAAACGTTACCTGTAAACAGCACTGCCTATCCCGATCAGTTAATATTCAGTCAACGG 3536  
3506 GGAATTTACAAAGCACACAGATTATGCTTTACTGCTCCACAGAGCCCGCAGTCCACAG 3565  
3537 CACCCCTCCGATCTGAATAGATAGTATTCGCAAGTTTAAATGGTTGAGCCTAGAA 3596  
3566 TCTACGCCCGTGGCGAGTCAAAATAGTGTGGCTAGGTTAAGTGAATTACAGCCCTGGA 3625  
3597 ACTGCGCCGAGGAGAACCGGTTGCGGTGGCGGTAGCCAGGGGAGACGGATGGCGCG 3656  
3626 ACCGAGCCATCGAGATGTCGCCACCAAGAGAG--GCCAGCGGAGATGATGCTGG 3682  
3657 CTTGAGGACGAGATGCTCTGCTGCTTTATCAGCGTCCGCTTGTCTATAGCCTTTCTC 3716  
3683 CGTAAGCACCGAGGTTTCTGCTGCTTTATGAGCGGAGTTCTGTCAATGGCCTTGGCC 3742  
3717 CCACAGCATATAAAGCAGCACTCTGAAGCTTGGCAGAAAGCAGCTTATATGATGCTGG 3776  
3743 CCACAGGAGAAAGACGCAATGTTCTAATTTGAGGATATGCTACTGATGATGAATCG 3802  
3777 TGAACCAATCCAGCCAGGTTGGGTGTGAGCCCGCTTCCCTCTTCCACCGTCAAG 3836  
3803 TGAACCAATCCAGCCAGGTTGCTCTGAGCGCGGCTTCCCTTCTCCACCGTCAAA 3862  
3837 TGGTGTAAAAAGCAGAACTCCCGCTTTTGGCCCAACAGCTTAAGCAGCCGTTGAGAGGAG 3896  
3863 CTGTGTTTATGAGCAACACACCTTGGCGAGCCAGCTGTGAGGACCCCGTGGGAGGAG 3922  
3897 CGGTGACCCAGTACAGTATTTGGCTATCTCTTTAAATATTTTGAAGGCTTGGAGSTA 3956  
3923 TACTGAAATTTGGGAGCGGAGGCTCTAGCTCTCTAAAGATGCTTCTCAAACTGGGTGAA 3982  
3957 TACTGTAGTCCGAGCAACGCTAAAGGCTAGTCCCGTGGCTTGAACCGGATGTTAAGGAT 4016  
3983 CTGTGATTTGGATCCACACTTAACGCGAGCCAGTAGCTTGGCTTGTGGTACGGT 34042  
4017 CTGGCCAAAGATACAGACTTTAATATCTCTCGGTGTCGCAACAATAGCACCAATAA 4076  
34043 CTGGCCCTAAGATCAACACTTTAATATCTCTGGATCGACAGTGGGACCAACACATCA 34102  
4077 TTCTATCTATCGTGGATAAATCACGCTTACGTTGCTCATATCCATTAACGCGCTTCA 4136  
34103 GCTTGTCTGTGGGGATACACTGTGGTGTAGCTTAAGTTCCGATCTGTCTGAGCA 34162  
4137 GAGCCGCGAGCTTCTGTTTAAATGAGAAAGATTTAAATATCAAGCA 4188  
34163 GCGAGAGCAGTTCTGTTTTCAGAAATGATGAGGCTCAGAAAGGAATCCA 34214

## RESULT 3

US-09-230-371A-18  
; Sequence 18, Application US/09230371A  
; Patent No. 6348586  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A  
; APPLICANT: Russo, James J  
; APPLICANT: Edelman, Isidore S  
; APPLICANT: Moore, Patrick S  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND  
; FILE OF INVENTION: USES THEREOF  
; FILE REFERENCE: 45185-G-PCT-US  
; CURRENT APPLICATION NUMBER: US/09/230,371A  
; CURRENT FILING DATE: 1999-11-17

; PRIOR APPLICATION NUMBER: PCT/US97/13346  
; PRIOR FILING DATE: 1997-07-22  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 35100  
; TYPE: DNA  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-230-371A-18

Query Match 11.0%; Score 1002; DB 4; Length 35100;

Best Local Similarity 54.5%; Pred. No. 5.1e-292; Indels 253; Gaps 10;  
Matches 2414; Conservative 0; Mismatches 1765;

QY 1 ATGAGAGAGTTCCGTGGATGACCAACACACGTGCAACCAATCCGGGTTTCATCTGAAC 60  
DB 29792 ATGGACAGCTCGGAAGGGTGCACTGATATGACGAGCCAGCCCGGTTTCATCTCAAC 29851  
QY 61 ATGAGCTCGGATGCCAAGTCAGGGGTGCTGTGATCAGCTCAGTCCGCTGTCATAATA 120  
DB 29852 ATGACATCGGATGCCAAGTCAGGAGCGTAGTGAACAGATTGACAGTTGTCAAATATC 29911  
QY 121 ACTACACGCCCAACCGGAATGGTTGGTACGACTGCGCTTCGATCCGGCTGGAAGATCC 180  
DB 29912 ACTACTCGCCCGGAGATGGCTGTATGACCTAGATTCGATCCACTGGAAGACGAA 29971  
QY 181 GGGCCGTTCTTGGCGTTTACCGTTTATCTAATTAAGGGAACACTGCTGTGTGGGAAAGT 240  
DB 29972 GGGCCCTTCTGCGCTTTTCGGCATAGTATAAGGGGACTGACAGAGCGGGGAAAGC 30031  
QY 241 ACCAGCATATCGGCGCTGTACCAAAATTTAACTGCTGTATCAGCGGGGCGACCAATTA 300  
DB 30032 ACCAGCATATCGGCGCTGTACCAAAATTTAACTGCTGTATCAGCGGGGCTACAGTGGTA 30091  
QY 301 GCGGCACAGAACCTATCGCGTCGCTAAAGACGTTGTGCCAGCATCTTCAGCGCTTTT 360  
DB 30092 GCGGCACAGATCTTTCCAGGCTTTAAAGTCTTCTGTCTCCACTATATACCGCTTC 30151  
QY 361 GCGTTTAAGAGCCGACACATCAATATAGCCGTGAGAAAGCTCATCAGACGGAGCGGTA 420  
DB 30152 GGAATTAAGAGACACATTAATATCTGCGAGAGGAAAGTGCCCAAGTAATCTAGTCC 30211  
QY 421 TCCATAGAGCAATTTACGACACAGAGCTATCGAAGTATTGGCGGTTATAGTGGACATT 480  
DB 30212 TCCATCGAGCAATCTCCAGAGATACGAGCTGTGCTAGTCTGCGCACTGTACCGATAT 30271  
QY 481 ATGAAAGAGTTATGGCGAATAAAACCCCAATGGCATGTACGGGACTATATCCACGGCAAT 540  
DB 30272 ATTCAGAAATTTATGCGCAAGAAACAAAAGGGGCGAGTATAGTCCCTCTCTCAAAGCGCT 30331  
QY 541 TTTGAAACCTCTCGAAGATGACCGGACCGTGTGTTATGGACTTCCCAATATTTATTGTAATC 600  
DB 30332 TTCAGACTCTTTGCGGTATGGGTGGAGCCAAATTTGGAGAGTAAACATTTACGTGATA 30391  
QY 601 GACGAGCGGAAACCTGTCTCTTACATACCTACCAACCGTGTGTTCTTTTCTGTTTCTGTTTC 660  
DB 30392 GACGAGCTGGAACCTCTCTGCTCCCATATTTTGAACCGCGGTGTGTTCTTCTATTGTTT 30451  
QY 661 CTAAACAGCTGGCTAAATACCCCTCTTTTACCGCCAGGGGCGGTTCGCTCATAGTATGC 720  
DB 30452 TACACAGTTGGCTGGACACCCCGCTATACAGAAATGGTGTGCGCTGCTGCTAGTCTGC 30511  
QY 721 GTCGGTTCGCAACGACGACAAACCGCTTCCAGTCAACTTACAAACGAGGAGCGGAAAG 780  
DB 30512 GTGGGTCTCCACCCAGACGAGCGCTTTTCAACGAGTCTTCAACACACGACGAGCAGAGA 30571  
QY 781 ACGGAGATATCGTGTGCGGAGAACATCTAACTTATGATGAGGAGAGGTCGTATCT 840  
DB 30572 AACGAGATATCGCTGTGATATGCTCACTTCTTATTTGGGAAACCTGAGGTGCA 30631  
QY 841 GAGTACGTACACTTTGGAGAGAAACCTGGCGCTGTTTATAAACCAACAGCGCTCCACCGAT 900  
DB 30632 GATTATATTAGCTGGAGAGAAATTTGGGCGCTTATTTATAAACCAATAGCGCTGTACGGAT 30691

QY 901 CTCAGTTCGGACACCTTCTAATAATTTAGAGTAAATCTTCCATTCTCGACGAAGTC 960  
Db 30692 CCCAGTTTGGTCACATTGCTGAAGACCTTAGAATAATACTAGACATATCACAGAGTTA 30751  
QY 961 ATGAGTTACGTAGACAGATTGTGCTTCTTAAAGTAAGATTATGGATCCTTTAGAAATAC 1020  
Db 30752 ATGACATATATAGATAGTTTGTGCTCCGAAGTAAGATTCTGACCCGCTCGAGTAT 30811  
QY 1021 ATTGGCTGGACCGACTCTTTTGTACATAGCGAGTAAGGCGTATCTACAAACCTA 1080  
Db 30812 GCAGGTGGACAAGACTCTTCACTCACACGAGGTAAGGCTTTCTTGGCAAGCTG 30871  
QY 1081 CACACATGTCTAACTAGGGGCGATACACAGGACACGAAGCTCTTTACCTGTCCCGTG 1140  
Db 30872 CACACCTGCTGTCGAGTATAGGATGCTGTCTCCACAAAGCTTTTCACTGCCAGTG 30931  
QY 1141 GTGTGAGAGTGTGTGAGACCGTTTGAGGAATACAGGGGCGCTCAACCTCACCAAC 1200  
Db 30932 GTCTGTGAGTGTATACAGAGCCATTTGAGGAGTACAAAGCGGCGTAGCGCTCACAC 30991  
QY 1201 CTCACGTGACCGAGTGGGTGACAAAACCTCTTTAAGTTAAGTAAATTTATCGCAGTTT 1260  
Db 30992 ATGACTCCCATAGATGGGTAAACAAAATCTTTTCAAGCTAAGTAACTACTCGCAGTTT 31051  
QY 1261 GTGACACGACATGTCCATAGTGGCAACGGAATCCACTGAAGCTTGGACCCAGTTACG 1320  
Db 31052 GCTGATCAGACATGGCTGTGGTGGACCTATATACACAGCGCTCCACAGATCACG 31111  
QY 1321 TTTATACCAAGTTTGTAAACACAGCCAGCTGTCTTTAATCGAAAAACAAAAAATGC 1380  
Db 31112 TTGCCACTAAATTTGTCAAAACAGCTATGTACCTTACTGGAAGAACAAAAATGT 31171  
QY 1381 ATATGTGGTTCAGGGTACGTACTTTGAGTCAAAAGATCTCTAGACAGCGAATCTCTTC 1440  
Db 31172 ATATGGGTTTCAAGGTCATACCAAAAGATTCAAGTCCATCTAGACGGGAGCTATT 31231  
QY 1441 GTGAGAGCGATTCGAGACCGCTCGGAATAGTATACGGTTTCTTAACACGCTACTG 1500  
Db 31232 ATCGAAAGTCAATCGCAGCAACCCCGCTTATGTGTACAGTTTCTTTAGTACCTTGCTA 31291  
QY 1501 TACAACGCCATGTACTCGTTTCAAGCTACGGGTGACTAGGTCAACAGAGAAATATCTG 1560  
Db 31292 TATTAATGCCATGTACTCATTTAGCGCACGGGTGAAGCAGGGCATGAAGATTCCTC 31351  
QY 1561 CAAGACTTAAATTTGACCCCTCCCGGCGCTCTGGCAACCGGGCGTGTAGACCTTCAA 1620  
Db 31352 AGGACCTCAGGGAAGTCCCGGTGTCTCAAGAGCTGATCTCTGAGATGAGCTCCGAGGAC 31411  
QY 1621 ACGGTTCGTGAAGATTAAACCTGGAAGACGACATCTTACACAGTGTGCACTCCCGG 1680  
Db 31412 GTTCTGGGCGAGGAG-----GGGGACACAGATGCTTCTACCTCCGCGCAGCTCCCA 31465  
QY 1681 CCACCCGCGGTATCACCTCCCTCCAGTTTGGTGCACAGTACTGCGCCCTAAAGGAC 1740  
Db 31466 CCATCCCCACCCACCGGCTCTTCCAACTGTTGGCTTATTACTCCGGGGCAAGGA 31525  
QY 1741 GTGTTGCTCCAGAAATAAGGTGGCTGTGCTGGTTTGGCGGGAGTTTGAGAAGGA 1800  
Db 31526 CTATTCGCAACAGGCTGGCCCTGGACGCGGACACTTTGGTACGAGTTCTCCACTCC 31585  
QY 1801 ACGTTTCCGGGTTTACGGTTTAAATGTCGTGAAGGACGAGTTGACTTGTCTCCCT 1860  
Db 31586 GATTTTCAAGTTTACGGTGAACATCGTGTGGAGATGCGTGGAGTTTGTGTCACAT 31645  
QY 1861 TCAGAACCTCTCAACGGGCTGTGGCGTTTGCATCGACCGTTGAATCGTATAAATTAAG 1920  
Db 31646 TCCCCCGGCTCCAGGCTCTAGTGGCATACGATCCACTATAGACACTATATAATCCAG 31705  
QY 1921 GGGTACAGTTTATACGGTAGCGTTTCGGTCTGTGAGGTTTGGCACTCAGTGATGAC 1980  
Db 31706 GGAATATGTTCTCCAGTGAGATTCCGCGCTCCAGGAGACAGCGCTCAGCGAGGAC 31765

QY 1981 CTCAGGAAGAATGCCCCCTCCTCGTGTGAGGACTCTAGCGGTTTTTATCGCGTGCTTA 2040  
Db 31766 CTGCGCAGAAAGATGCCCCCTCCATAGTTGTCTCAGGACTCATCGGGTTTCATTTGCCCTGCTG 31825  
QY 2041 GAGATAACATAACCAATTTGACCGAAACCATGAGGACGAGCGTTTTCACAGTGTGC 2100  
Db 31826 GAAATAACGTCACCAAGATGACAGAGACCTCCAGAGTGGCGAGTGTGTTTAAATATGT 31885  
QY 2101 TGTGCGGGGACTATGGGGTCACTCAAAATTTAGCCATGACCATCGTAAAGGACAGGGA 2160  
Db 31886 TGTGCGGGGACTACGGTATCAGTTCTAACTCTGGCTATGACCATAGTAGAAGCAGAGGG 31945  
QY 2161 ATGTCTGTGGAGCAGTAGCGGTAGTATTTGGGTCCCAAGAGAGCTCCAGACAGCCAC 2220  
Db 31946 GTTTCACTAAGTAGGGTGGCCATATCGTGTGGCAACCAACCCCAATATCAGAGCCAGTCTA 32005  
QY 2221 GTGTATGTAGCAATATCAAGGGCTGTTAACTCAAAATTTATTTGGTCATGGACAGCAACCCC 2280  
Db 32006 GTGTATGTGGGTGTATCCAGGGCCATCGACGCTCGTTACCTTGTAAATGGACAGTAATCCC 32065  
QY 2281 CTTAAAAACCTCTCTCAGAGAACAGTGTGATTAACCTCCGCGCAAGCATATAGTCCGCGCC 2340  
Db 32066 CTTAAGCTAATGGACCGGGTGAAGCCGAGTCCCGCATCTCTCAAAAGTACATCATCAAGCC 32125  
QY 2341 CTCACACACCAAAACACAAACCTCATCTACTAA----- 2373  
Db 32126 CTATGCAACCCCAAGACTACTCTGATCTACTGACCCGTAACCCCTCTCTTAGACACTGAT 32185  
QY 2374 -----ATAAATAAAACAAAGCAAAATGGAAAAATTTG 2406  
Db 32186 GTGTTTGGGAATAAAGCATGAGACTTGACACTATAATGGTCTGTATTTGACACCAATCTT 32245  
QY 2407 TTTTATTTTCACTCMAACACCGCCACGGATAGTTGTCAATTTCCACACACCGGGGGTGG 2466  
Db 32246 TTTATTTACAGTCCAGCCGCGAGTTATATGCAACGTTTCCACACAGGGGGTGGCGTGG 32305  
QY 2467 CGGCAATAGTTTGAAGCAGCAAGAAACGCGCGGGCCCTCTGGGTTAGAAATGAGCCCC- 2521  
Db 32306 AGGCCAAGGATCGGGGTGGTGGTGCCTGACCTGGACCCCGGTAGTTGTGTTCTCGATG 32365  
QY 2522 -----GCCCCCGGTGTCAGGTCATACGTCCTTAGGGCCCTGGT 2564  
Db 32366 AAATCGAGTGGCGGAAGTACTGGGAGATGGGTTGGAGGTGACCCCTTTGTCTCGAGG 32425  
QY 2565 GTACCAACATAAATCAGCTGTCACTCCGTTGAGTCCCTGCGTGAAGCGCGCTCTTTG 2624  
Db 32426 GAGACAGATCAGCTCAGCGCGGACGAGGGCTCTCTGCTGTGTCTCTCCCGAGGATA 32485  
QY 2625 CAGCGGCGACGAACACCGCGCTTGTAAAGTATCT-----GCAT 2664  
Db 32486 TAAATATCAGCGACCCACTGCTTTTGGCGCTTAAAGTTGGTGTCTCTGGACGCGACA 32545  
QY 2665 CGTCTGCTGTCTGAAGATGAACACAGGTCAATCAGATGAGCGCGCCCTGCGCTCAT 2724  
Db 32546 CATCTCGCTACAGAGGAGGGGTAGACTGCCCTTTTGGCTTCTGGGCCACGTCATGA 32605  
QY 2725 CGGCGTCTGATTTGTCCTCATGTTCAACAGAGCCTCGGTGCTCACTTCTGCTGCTCATCTG 2784  
Db 32606 GCCCGATTCTCTGACTCAATACTTCCCTTGTCTCTCTCCGCTCTCTCGGACGAGGGTG 32665  
QY 2785 AAATATATGTCTCTTGTCTGTCAGTTGATCCAGCAACACAGGTTCTTCTTCTCATCGG 2844  
Db 32666 GTGTGTGGGAAAAATGGCGCGCTCGGTAAACCGCGCTCATTTGTTCAGTCCGAGAGT 32725  
QY 2845 AAGACCATCCGCGCTCTGCGGTGATCTAGATATATACCTGACAGGGCGGCGAGGCC 2904  
Db 32726 TGGAACTGTCTCGCTATCAGAGTCCGATGTCAAGTCCAGTCCGCGTGGTGGCGCG 32785  
QY 2905 G----- 2905  
Db 32786 GCAGGGGCGCCACAGAGGCCCTTATCAGGGTCTGTGTATGGTGAACCTTTGTGTCCAG 32845  
QY 2906 -----TGGCAAAAGCCGCTCAGCTTCTCCGCGCACTTTTATCAGTGT 2946





7759 CAGTTGGCGCAGATTCTGCACGAGTCTACCGGAAATGAAGCCCAAGTGTGTTGGGGC 7818  
Db AGCCATATGTAAGTACTACACGAAATATACACTGAAATGAAGCCCAATGCTTGGGGC 2826  
QY GTGGCGTCTCTCATATGTTGAGGGCGCTATCATGCTATTGCCAGCTCGGTCTCGT 7878  
Db CTGGAGCGCAGTCTGCTGCAATCGAGGCCCATATGATATTAACCTCTTCCCTACTGAA 2886  
QY CACTTATATACCTGAGTCCGAGGGAATCTGAACTTAGACTCAAGACCTGTGCA 7938  
Db GTCTACACAGCTAGATACCGCGGGCTGCTCTCTGAGCAGTCCAGGGCCCTCTGCT 2946  
QY ACTTGTGTTTCAACCCATCTTTCTTCAAGAAATCTTAGCCCTATGAGCTACTGGCAA 7998  
Db TTTGGTTTCCAAACCGTCTACTCTTCGAGGATTAATGCGCGCTGGAGATCATGACCA 3006  
QY CGGAGATGTTCTGACATATTTTACCATCACCGTACCGGACCGGAGAGGGGCC 8058  
Db GGTCAGCTCGCCCTGAAATCTTTTACGATCATACCGGTTCTCTGAGAAACCGCGGC 3066  
QY TGTCTGACTGGAAGTACTAGCGGATGACGCTGTCCGGGAAGCAGCTTGTCCCGATTC 8118  
Db AATTACACCGCAAGTCACTGGACTGTCTATCCAGGAAGCGTCTCATGCCAGATC 3126  
QY CTTAATCTGCGAGTATGCGAACCGGGCTTCTCCGGCACCCCTGTTGACCTCAGTAA 8178  
Db TTTAAATTTGCAATCTCTGAGCCAGGACTGTTCGGGCTTCCATGTAGACCTCAGCA 3186  
QY TGTCTTAGAAATCAGAAATCATCTCAGCGGCCCAACCCCTGAGTCAATTTGTCAATCA 8238  
Db TGTCTGCGCAAAACCGCGTTATTTCTGAGCGCCCTGCGCTGAGCAGTTTGTCAATG 3246  
QY AAACAGCAGCCAGTCTGCTCAGTCACTGAGCAGTATTTAGCGCAACCGGCGGTGT 8298  
Db CAAACCCATCCAAATGCGCACACCGCTGAGTCACTATCCCTTTAACCC---CATCGG 3303  
QY TCCGCGCCATGTTTATGACACGCTGGAAGCGGTGTCAAGAGCATTCACACACAGGC 8358  
Db TACAGACCGCGTTTATAGTACGCTGCGCGCGGTCAAGAAATGTTGTACACAC 3363  
QY ACAGAC 8364  
Db ATCCAC 3369

## RESULT 5

US-08-757-669A-19  
; Sequence 19, Application US/08757669A  
; Patent No. 6183751  
; GENERAL INFORMATION:  
; APPLICANT: Chang, Yuan  
; APPLICANT: Bohenzky, Roy A.  
; APPLICANT: Russo, James J.  
; APPLICANT: Edelman, Isidore S.  
; APPLICANT: Moore, Patrick S.  
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/757,669A

; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 45185-F  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 35100 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-757-669A-19

Query Match 7.9%; Score 714; DB 3; Length 35100;  
Best Local Similarity 54.4%; Pred. No. 1.3e-204;  
Matches 1569; Conservative 0; Mismatches 1280; Indels 37; Gaps 5;

QY 5511 GGCAAGGCTCATTTGCCATTTCCAGATTTGTTAAATCTGTTGAAACATACGAGTCTAAAT 5570  
Db GACACTGGCAATGATGACTCCACCGTCTCACGTCGAGGTAAATGTCTTATGAAT 548  
QY 5571 CAAGTATACCAACGAAAGTTCAACAGAGAGCAACGCGCTGTTTCCACCCGAGCATC 5630  
Db GTGGTAGAGCCAGCAGAGCTTCAACACAGATGAATCAAGCTAAGCTCCGCGCGGAATC 608  
QY 5631 CCGCTCACAAACCCGAACTCTCTAGGAACGCATCAAAAGCTTCAGACAAATTCGA 5690  
Db CACATCCAAAGGGGTATTGCTCCGCTGTCTGTTAGTCTGGAATAGAACTCAGA 668  
QY 5691 AACGAAACCCCTGTTCTGAGCACCACTCTGTTTAAATCAATGCGGTTAAAGAGAGCGGC 5750  
Db AAAAGACACTGACCCACCAAGAGAACCTGCGCTCTTGAAGTTGATGAGCCCGCAGA 728  
QY 5751 CACACACTGCTGTTGAGGATTACGTTAGTGGCTTGCAGCTTGTCTGACAGCCCG 5810  
Db AAGATGTTCTCCGTTGGACAAAGAGCTTGGGGGGGAGAGATGGCGCTACAGTGGT 788  
QY 5811 TAGCTGCGTGGCGCGCTCGCGACCTTTGACAGCCGCGAGCCAGTTCAGAGCGAGGCA 5870  
Db GATTTCTTACCAACGGTCTACATTTGGTGGGACCCACAGGCGCTGTTCAGTATCAGCAT 848  
QY 5871 GAAGCGCTCTTTGAGTCGCGCATATCTCGGCATAGGAGATTCGCTCGCGTATCTTT 5930  
Db AATCTATCTTTGAGTCTATCCAGATCAAGTCAATGTCAGATGCTGTTGCGCTGCAATTT 908  
QY 5931 TAGTGATATATATCTGTTGGGTAAACAACTCACTCCCTGTAAAGGGGAATAAGGTC 5990  
Db TGCCCGCATGTACATTTCTCTCCCATATTTTAAACATCTGTAATCTGGAAGTAGATT 968  
QY 5991 GGTCTTTCTATCGTTTCTCTGAATCTATATAGTATACATCTTTAAGTATTAACAAAGACGC 6050  
Db CAGTCTGGTGTGAGCCCCCGGGGAGCCAGCGTATGCTTCAGGACCCAGGAGCGC 1028  
QY 6051 TAAAAACCCAGATGTCGACACTTTTAAAAATCCGAAAGCGAGCGGATGCTCTATC 6110  
Db TAAGAACCCCGGTGTCCGCGCTCCGGAACAGACCTCTGGAATACGCTCGTCTGTGAC 1088  
QY 6111 CGCGTTTCCATGTCTTCAAGGTAGTTCAGATGAGATCTCCAGTATAGACTCTGTCTCCCT 6170  
Db GAAACCCGATGTGTACCGAATGCCAATCTGTGCGCTCCAGCTCTCAAAATTTTCAATC 1148  
QY 6171 GTTAAACCCCTTAAGTGAATGAAACAGCCATGACCGTGAACCTTCTGATGCTCT 6230  
Db TCCAAATACCCGGAATTTGGGATACACACCTCCATGTTTCACTACATGATCAGTGGTCTC 1208  
QY 6231 AAAAATAACCCCGAAACAGAGCTTAAATACCAATGACTGTACCCCTTACCCCAAGCCAGC 6290  
Db CCCACCCCAACCCCATAGGACCCAGCTACAGCTTATCTCTCCACTAAATACCAGGACGTA 1268

QY 6291 CCCGTAATTT-----AGACGAGGTCAGTAAACACGCTATCT 6330  
Db 1269 CCGGGACTATTAAGCCGCCCCAGAAACAGTAGCTGGTGGCAATGACAGCTCCCT 1328  
QY 6331 TTAATAAACCCATACGTGGAGTTGTAAAGTAAACAGCT-CGTGTATATTTTACAGCGCTG 6389  
Db 1329 TTAATAAAGTCAACCTTACTCCGCAAGGGTAGTCTGTGTGAGAACTACTGTCCAGCAGC 1388  
QY 6390 TCATAAATGATACCGACGACAAATCAGGTAATTAATAATTTTATTTTATTCAGATTCAGGA 6449  
Db 1389 CACAAATATGCGCAAGATGACAAAGTAAAGATCGACC--TTTATTTATTAAGTAACTGAACA 1446  
QY 6450 GGGCGTGTATACGTGTAGTGGAGCATATCTCGCAATCTCGCAATGGAATGGCTGATGAAT 6509  
Db 1447 ATGGGTGTATTAATATGTTAGTGGAGCAGAGTTTCGCCAAGCTCTAAGTCCGAACAGT 1506  
QY 6510 CCACACTTATGTCGCGATCAGGTGCTCTGTGTACTTTTACTAGTCCGATGACCTTAAT 6569  
Db 1507 CGGGTGTACAGGCTCTTATTAAGTGTTCGGTGTACTTGACCAAGCCGCGAACCCTAGGT 1566  
QY 6570 TTGGCGTGTATGAGTGTGTAACAACCAAACTGCGTGGCTCGCTTTTATAGAGCTCTA 6629  
Db 1567 TGGGTCTGTACAGGTCTGTACAGCAAAAGAGATCGGCGGTGCTTTTACAGAGAGTTA 1626  
QY 6630 TGCAAGCTTAAATCGTCTCAATAGCTGTCTTTAAATAAGCACTCTGTACATCTTTA 6689  
Db 1627 GGGACGTGCTGATTAATGTGACAAAGCTTCTGCTGAATGCAACCCGCTGTACATCTGAA 1686  
QY 6690 TGACAGTGTGTCNAAAAAAACAAGATTTAAACACAGTTAAACTGCTGCTCGTAGT 6749  
Db 1687 CGACAGCTGTCAAAAAAACAAGGTTCAAGTGCACGTTAAATCTGTATCTCGAAAGT 1746  
QY 6750 CCTCGTATATAACCTCTCAACAGAGAAATTTTAAACCAACCTCGCAGGACTGAA 6809  
Db 1747 CCTGTAAATGACAGTTTCTACCAAGAAACTTTTACACGCTGGCCATCCACTGAA 1806  
QY 6810 ACAGATGTAACGACAGTCTGTGAAAGTGTCTATCATCTCTTCACTCCCTCCCTTTGAAAG 6869  
Db 1807 AGAGGGAGCACAGCTCCGCTGTGGCGTTGTAGGATATCCCTAACTTCGGAGCGGAGAC 1866  
QY 6870 TTTTGTGTACACGACCATGTGATAGTAGGACACAGCTCAGGGGGAGTCCGATCGTAG 6929  
Db 1867 GCGCGAGCTCCCAAAATGAGAGGAGCACCACTCTGTGCAGTCCGCGTCTGGGTT 1926  
QY 6930 GGGGTAACAATTCGTGGCGGTGGGGTATCCAGTTTTAGATCTCCGCGACGACGCTA 6989  
Db 1927 CTGATTCAGGGGCGCGTGTGGGGTATTGGAGAGTCAAACTCTGGGCACTCCCTTAA 1986  
QY 6990 CGAGTCTTCGTCAAAATCTACGACGCCGCATCCACCGGGCAGTACTGAGTTAATAA 7049  
Db 1987 TGAGCTCTCTCAAAACCTATGACGACGCTCCACTAGTGGCAGCATGCCGTTAATAA 2046  
QY 7050 GTCCCCGAATTTCTATCGTTCGTAAATTTGTGGAGCTCCTTTAGACAGTAAACCATGTTGC 7109  
Db 2047 CACCCCTTATCTGTCTGTGCAAGTTGTACAACTGCTGAGGGAATAAGCCAAATTCG 2106  
QY 7110 TCCGACACTCGGTACGATGAACGCTCTCTGTCTGTGTGTCATGACACCTCGAGTA 7169  
Db 2107 CCCTAGCCGCGGAACCAAGTACGCTCTGTGTGTGTCGTGAGCAATATCTGAATGG 2166  
QY 7170 TCTTAGACAGTATAGGGCCCGTTTCCACTCTAAGGGTGTGCGCTGTGTAGAGCGGATTGC 7229  
Db 2167 TCTTTGCAAGGTAAGGGTCTTCTCAACGTTTAGAGCGGGTACGTGGCAGTCTGGATTGA 2226  
QY 7230 GGGCATTTGACGGTCAGTGTGTCTATTTACCGGACTATCATCGGAGTGGCGGATGATGGA 7289  
Db 2227 GGGTGGGACGAGCAGGGTATCTAACTCTCTGAAGTATCTGATCCCAAGGACGGGTAAATGAT 2286  
QY 7290 ACTTGTACATATGATTTAAAGATGGTCTTTTCCCGTAAAGGGCACTCATGGTCTGTGTTT 7349  
Db 2287 ACCTAAACAGATGCTGAACAGGTGATCTTTAAGGGCGCTTCTCGATGTCTATGTTAAAAA 2346

QY 7350 A-----ACGTTCCAGGGTGTAAAAAACTCAGAAAGTCTCGTCTGGCAACATT 7398  
Db 2347 CTATGACACGCCACTCTCTCTCTTAGGGTAAAGAGCTTCGGCGGTCTGTGTGGAAAGCTT 2406  
QY 7399 TTGGGACTCTCCGAAACCCTCAGGGATTAAGCTGTACCAACGGGGTCTCTTTATATAAAGA 7458  
Db 2407 CGTGGCGCTCTCGGACGAACCTGAAGCCCAACTCTACCAAGTGTGTCTCTTTATAAATGA 2466  
QY 7459 AGCATAGCTCAAAAAACATGACCCACGTCAGCTAAATAGCTCACAGAAAGTCTGTGCG 7518  
Db 2467 CGCATAGCAAAACATCTCATCCAGTGACCTTAATAGAGTGGTGAAGATGTGTGCAT 2526  
QY 7519 ATGTATTTAATGAAGATGTCCCTCTGGTCCCATCTCCGGGCTCATCCGCGATCTGAA 7578  
Db 2527 TCGAATTAAGAAAGATGTTCCAAGCTTGGTGGCTATGTGTCTGTGTGTACAGACATTAA 2586  
QY 7579 TCTCTTTAACTGCTCTCTGCTCTACCGGGTTCCTGGGTAAAAACCGCGGGGCGCCAC 7638  
Db 2587 CATGTTTAACTTCTCTCTCTTTTCTGTCCTCTCGATGAGGACCAAGGCGCGCGCGG 2646  
QY 7639 CTGTAAAGTCCCTGCGCAGAAATGCGCCACAGGGCATCTGTAGAAATCTCAACGAGAGGC 7698  
Db 2647 GTACAAAGTCCATGCGCAGAGGCAATCCCAAGGCATTAATTCGATCTCTCACGAGAGGAT 2706  
QY 7699 CTTTGTGTGACGAAAGATGTTCATAGCTGTCTGTCAGCGGGTCTGTATACCGCC 7758  
Db 2707 CTTTATTTGCAAGAAAGAGCATTTCTGACGCGCATGACGCGGGTGTGAGCTGCTCC 2766  
QY 7759 CCAGTTGGCCAGAGTTCTGACGAGCTCTACGAGAGTCTACCGGAAATGAAGGCCAAAGTGTGGGGC 7818  
Db 2767 AGCCATATGTAAGCTACTACAGAAATATACACTGAATGAAGGCCAAATGCTGGGGC 2826  
QY 7819 GTGGCTGTCTCATATGTTGACGGCGGCTATCATGGCTATTGCGCATCTGGTCTCTGT 7878  
Db 2827 CTGGAGGAGTCTGTCTGCAATCGAGGCGCATTAATGATTAATTAATCTCTTCCCTACTGAA 2886  
QY 7879 CACTTATATACCTCGATGCGCAGGAGAAATGGAACCTTAGACTCAAGACACTGTGCA 7938  
Db 2887 GCTCTAACAACAGCTAGATACCGCGGGTGTCTCTGAGCAGTCTCAGGGCGCTCTGCT 2946  
QY 7939 ACTTGTCTTCAACCCATCTTTCTCAAGAAATCTTAGCCCTATGCGAGTACTGCGCAA 7998  
Db 2947 TTTGGTTTCCAAACCGTCTACCTTCCGAGGATTAAGCGCGCTGGAGATCATGACCAA 3006  
QY 7999 CGGGAAGATGTTCTGCAAAATTAATTTTACATACCGGTACGGCCGAGAGAGCGGCC 8058  
Db 3007 GGGTCACTCGCCCTTGAAACCTTTTACAGCATCACCGTCTCTGTGAGAAACGCGCGCC 3066  
QY 8059 TGTGCTGACTGGAAGTACTAGCGGATGACGTCTCCGGAAGCAGCTTGTCCCGATTC 8118  
Db 3067 AATTACACCGGCAAGTCACTGGACTCTCTATCCAGGAAGCGTCTCATGCCAGATC 3126  
QY 8119 CTTAATCTCGCAGTATGCAACCGGGGCTTCTCCGGGACCCCTGTGTGACTCAGTAA 8178  
Db 3127 TTTAATTTGCAATCTCGAGCCAGGACTGTGCGGCTTCCATGTGTAGACTCTCAGCGA 3186  
QY 8179 TGTCTAGAAATTCAGAAATCATCTCAGCGCCCAACCCCTGTGACTCAATTTGTCTATC 8238  
Db 3187 TGTGCTGGCAAAACCGCGCTTATTTCTGAGCGCCCTGCGCTGTAGCCAGTGTGTCTAG 3246  
QY 8239 AAACACGACCCCACTGTCTGCTCAGTCACTAGCAATTAATAGCCAAACCGAGGGCGTTGT 8298  
Db 3247 CAACCCCATCCCAACATGCCGACACCGCTCAGCATCATCCCTTTAAACC---CATCGG 3303  
QY 8299 TCCCGGCCAATGTTTTATGGACACGTTGAAGAGCGGTGTACAGAGCATTCACACACCGC 8358  
Db 3304 TACAGACCGCGGTTTTATAGTACGTGGCAGCGCCGCTCACAGATATGGTGTACACAC 3363  
QY 8359 ACAGAC 8364  
Db 3364 ATCCAC 3369

RESULT 6  
 US-09-230-371A-19  
 ; Sequence 19, Application US/09230371A  
 ; Patent No. 6348586  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chang, Yuan  
 ; APPLICANT: Bohenzky, Roy A  
 ; APPLICANT: Russo, James J  
 ; APPLICANT: Edelman, Isidore S  
 ; APPLICANT: Moore, Patrick S  
 ; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND  
 ; TITLE OF INVENTION: US/09-230,371A  
 ; FILE REFERENCE: 45185-G-PCT-US  
 ; CURRENT APPLICATION NUMBER: US/09-11-17  
 ; CURRENT FILING DATE: 1999-11-17  
 ; PRIOR APPLICATION NUMBER: PCT/US97/13346  
 ; PRIOR FILING DATE: 1997-07-22  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 19  
 ; LENGTH: 35100  
 ; TYPE: DNA  
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
 US-09-230-371A-19  
 Query Match 7.9%; Score 714; DB 4; Length 35100;  
 Best Local Similarity 54.4%; Pred. No. 1.3e-204;  
 Matches 1569; Conservative 0; Mismatches 1280; Indels 37; Gaps 5;  
 QY 5511 GGCACGGCTCATGGCCATTTCCAGATTGTTAAATCTGTTCGACATACGAGTCTTAAT 5570  
 DB 489 GACACTGGCCATGCATGACTCCAAACCGTCTCAAGTCCGAGGTAATGTCTCTATGAGAT 548  
 QY 5571 CAAGTATAACACACACAGCAAGTTCAACAGGAGGCAACCGCGCTGGTTCCACCGCGACATC 5630  
 DB 549 GTGGTAGAGCCAGCAGAGCTTTCAACACGATGAATCAAGTAAAGTCCCGCGGAATC 608  
 QY 5631 CCGCTCACAAACCCGAACTCTTAGAAACGACATCAAGAGCTTCAGACAAATTTCCGA 5690  
 DB 609 CACATCCACAAAGGGGTATTGCTCCGGTGTCTGTATTAGTCTGGAATAGAAATCTCAG 668  
 QY 5691 AACCGAACCCTAGTTCTTGAGCACCATTCTGGTTATAAATCTGCGGTTAAAGAGACGAGCGG 5750  
 DB 669 AAAGACACTGACCCACAGAGAGACCTGGCGTCTTGCAAAATGATGAGCCCCGAGA 728  
 QY 5751 CACACACTGTCTGTGTGAGGATTCAGCTTAGTGGCTTCGAGCTTCGACAGCCCG 5810  
 DB 729 AAGAAATGTCTCCCGTGGGCAAAAGAGCTTGCGGGGGCAGAGATGGCGCTACAGTGGGT 788  
 QY 5811 TAGCTGCTGGCGCGCTCGGACCTTTGACAGCCGACAGCTTCCAGATTCAGAGCGGAGCA 5870  
 DB 789 GAATTTCTTACACCGGTTCATACATTTGGTGGCACCACAGGCTTTTCAGATACAGCAT 848  
 QY 5871 GAAGCGCTCTTTGAGTCCGGCCATATCTCGGCCATAGGAGATTGCGTCCGCTGATCTTT 5930  
 DB 849 AATCTATCTTTGAGTCAATCCAGATCAAGTCAATGTCAGATGCTGTTCCCTGGCATTT 908  
 QY 5931 TAGGTGATATATATCTGTGGGTAAACAACTCACTCCCTGTAAAGGGGATAGATTC 5990  
 DB 909 TGCCCGCATGTACATTTCTGTCACATATTTTAAACATCTGTAATCTGGAAGTGAAT 968  
 QY 5991 GGTCTTTCTCTCTGTTCTCTGAATCTATAGTTACATCTTTAAGTATACCAAGACGC 6050  
 DB 969 CAGTCTGGTGTGAGCCCCCGGGGAAGCCAGCGTATGCTTCAGGACCCAGGAGCGC 1028  
 QY 6051 TAAAAACCCAGGATGGTGCACCTTTTAAAAATCGCAAGAGCCAGCGATGCTCTATC 6110  
 DB 1029 TAAGAACCCCGGGTGTCCGCGCTCCGGAACAGACCTCTGAGAAATACGCTCGGTCTTGAC 1088  
 QY 6111 CGCGTTTCCATGTCGTTCAAAAGGTAGTCACGATGGATCTCCAGTTAGACTCTGTCTCCCT 6170  
 DB 1089 GAACCCCGATGTGTACCGAATGCCAATCTGTGGCTCCAGCTCTCACAAATTTTCATC 1148

6171 GTTAAACACCCCTTAACCTGGAATGGAACAGCCATGACACACCGTGAACCTTCCTGATGTCCT 6230  
 DB 1149 TCCATATACCCCGGAATTTGGGATACACACTCTCAATGTTTCAGTCAATGATGATGGGTCTC 1208  
 QY 6231 AAAACTAAACCCCGGAACAGAGCTAAATACCAATGATGTCACCCCTTACCCCAAGCCAGC 6290  
 DB 1209 CCCACCCAAACCCCATAGGACCCAGCTACAGCTATCTCTCCATTAATACAGGACGCTA 1268  
 QY 6291 CCCCGTACTATT-----AGACCAGGTGAGTAAACCAACCGTATCTCT 6330  
 DB 1269 CCGCGCTACTCATTAAGCCCGCCAGAAACAGTAGTGGTGGCAATGACACGTCCTCT 1328  
 QY 6331 TTAAAAACCCCATAGTGGAGTTTGAAGTAAACAGCT-CGTGTATATTTTCAGACGCTG 6389  
 DB 1329 TTAATAAGTCAACCTTACTCCGAAGGGGTAGTCTGTGTGAGAAATCTCTTCCAGGAGC 1388  
 QY 6390 TCATAAATGATACCCAGCAGCAATCAGTAAATTAATTTTATTTTAAAGATTCAGGA 6449  
 DB 1389 CACAAAATGGCGCAAGATGACAGGTAAGATCGACC--TTTTTATGTTATCTGAACA 1446  
 QY 6450 GGGCGTGTTTACAGTGGTGTAGTGGAGCATATCTCCCAATGGGAATGGCTGATGAT 6509  
 DB 1447 ATGCGTGTTTACAAATGCTAGTGGAGCAGAGTTCGCCAAGCTCTACGTCCGAAACAGT 1506  
 QY 6510 CCACACTAGTGTCCGATCAGGTGCTCTGTACTTTTACTAGTCCGATGAACCTTAAT 6569  
 DB 1507 CGGTGTGAGGCTCTTATTAAAGTGTTCGGTGTACTTGCACCAAGCCCGGAACTTAGT 1566  
 QY 6570 TTGGCGTGTTTAGTGTCTGACCAACCAAAACTCCCTGGCTCCGCTTTTATAGAGCTCTA 6629  
 DB 1567 TGGGTCTGTACAGTGTCTGACCAAGGATTCGGCGGTCTTTTTCAGGAGAGTTA 1626  
 QY 6630 TGCAGCTTTTAACTGTCTCAATAGTGTCTTTTAAATGCAACTCTGGTACATCTTTA 6689  
 DB 1627 GGGAGCTGTCTGATTTATGGAAGCTTCTGCTGTAATGACCGCTGTGTACTCTGAA 1686  
 QY 6690 TGACAGTGTCTCAAAAAAACAAGATTTTAAAAACAGTTTAAACTGCGTGTCTGGTGT 6749  
 DB 1687 CGACAGCTGTCTCAAAAAAACAAGTTCAGTGCAGCTTAAATCTGTATCTCTGAAAGT 1746  
 QY 6750 CCGTGTATATTAACCTCTCAACAAAGAAATTTTAAACCAACTCCGAGTACTGAA 6809  
 DB 1747 CCGTGTATATGACAGTCTTCTACCAAGAAATTTTACCAAGTCCGCTCCATCCTGAA 1806  
 QY 6810 ACAGTGTGTACGACAGGTCGTGAAAGTGTCTATCTCTTCACTCTCCCTCCCTTTGAAG 6869  
 DB 1807 AGGAGGAGCAGACAGTCCCGTGTGGTGTGTAGGATATCCCTAACCTTCGAGCGGAGAC 1866  
 QY 6870 TTTTGGTTTACACCGACCATGTGAGATAGGACACAGCTCAGGGGAGGTCCGATCGTGA 6929  
 DB 1867 GSCCGGACGCTCCACAAATATGGAGAGGACCACTCTGTGACGTCCCGTGTGGGTT 1926  
 QY 6930 GGGGTAAATTCGTGGCGTGGGGTACTCCAGTTTGTAGTCTCCGCGCAGGACCGCTA 6989  
 DB 1927 CTGATTCAGGGGCGCGTGTGGGGTATTTGGAGTCAAACTCTGGGCTCCCTTAA 1986  
 QY 6990 CGAGTCTCTCTCAATCTTACGAGCCCGCATCCACAGGGGAGTACTGAGTTAAATA 7049  
 DB 1987 TGAGTCTCTCTCAAAACCTATGACAGCGCTCCACTAGTGGCAGCATGCGCTTAAATA 2046  
 QY 7050 GTCCCGGATTTCTATCTGTTTGTGAGTCTCTTTAGACAGTAAACATGTTGTC 7109  
 DB 2047 CACCCCTTATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2106  
 QY 7110 TCCGACACTGCGGTACGATGAACCGCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7169  
 DB 2107 CCTAGCCCGGGAACAGGTACGGCTCGCTTGTGGTGTGGACCAATATCTGAATGG 2166  
 QY 7170 TTTTACAGGATATAGGCGCGCTTCTCACTTAAAGGTTGCCGTTGTAGAGCGGATTCG 7229  
 DB 2167 TCTTTCAGGATATAGGCTCTTCTCAAGTTTGTAGAGCGGATACGTGGCAGTCTGTGATGA 2226  
 QY 7230 GGGCATTTGACGCTCAGTGTGTGTTTCTTTCACCGACTATCATCCGAGTGGCGGATAGTGA 7289



Db 2227 GGGTGGGAGGAGGAGGCTATCTAATCTCTGAAGTATCTGATCCGAGCGGTATGAT 2286  
Qy 7290 ACTTGATCATATGATTTAAAGATGCTTTCCCGTAATGGCGACTCATGGTCTGTTT 7349  
Db 2287 ACCTAAACAGATGGTGAACAGGTGATCTTTAAGGGGCTTCTCGATGTCAATTTGTA 2346  
Qy 7350 A-----ACGTTTACAGGGTGTAAATAAACTCAGAGGTCCTGCGTGACAACTT 7398  
Db 2347 CTATGACAGCGGCACTCTCTCTTAGGGTAAGAAGCTTCGGCGTCTGTGTGGAAGCTT 2406  
Qy 7399 TTGGGACTCTCCGAGCGCTCAGGATAGCTGACGCGGTGCTTTTAAACGA 7458  
Db 2407 GTCGCGCTCTCGGAGCAACTGAAGAGGCCAACTCTACCGGTGTGCTCTTATAATGA 2466  
Qy 7459 AGCATACGTCATAAAACATGACCCAGCTGACCTAAATAAGCTCACGAAGAAGCTCTGTCG 7518  
Db 2467 CGCATACGAAACAATCTACGATCCAGTACCTAAATAGAGTGGTGAAGATGTGCAAT 2526  
Qy 7519 ATGATTTGTAATGGAATGCTCTCTGCGTCCATCTCCGGCTCATGCGGATCTGAA 7578  
Db 2527 TCGGATTAAGAAAGATTTCCAGCTTGGTGGCTATGTGGTCTGTTTACAGACATTAA 2586  
Qy 7579 TCTCTTTAACTGTTCTCTCTACCGGGGTTCCCGGTTAAACCGCGGGGCCCCAC 7638  
Db 2587 CATGTTAACTTTTCTGTTTTCTGCTCTCGAATGAGGCAAGCGCGCGCG 2646  
Qy 7639 CTGTAACCTCGGTCGAGAAATGCGCCAGGAGCATCGTGAGATTTCTACCGAGGGC 7698  
Db 2647 GTACAACGTGCATGCGCAGAGGATCCCAAGGCAATTAATTCGATCCTCAGGAGGAT 2706  
Qy 7699 CTTTGTGTACCGAAAGATGTTATAGGCTCTGCTGCGAGCGGGTCTTTATACCGCC 7758  
Db 2707 CTTATCTGCAGAAAGGCAATTTCTGACGCGCATGCGGGGTGAGCTGCTCTC 2766  
Qy 7759 CGAGTGGCAGAGTTCTGACGAGCTCTACGCGGAAATGAAGCCAGTGTGTTGGGGC 7818  
Db 2767 AGCCATATGTAAGCTACTACAGAAATATACCTGAATGAAGGCCAAATGCTGGGGC 2826  
Qy 7819 GTGGCGTCTGCTATGTTTCCAGCGGCTTATCATGGCTATTGCGGACTCGGCTCTCT 7878  
Db 2827 CTGGAGGCGACTGCTGCTGCAATCGAGGCCAATATGATTAACCTCTTCCCTACTGAA 2886  
Qy 7879 CACTTATATACCTGAGTGGAGGAAACCTGGAACCTTAGACTCAAGAGCTGTGCAA 7938  
Db 2887 GCTCTACAAACGTAACGATACCGCGGCTGCTCTGTAGCAGTCCAGGCGCCCTGCTC 2946  
Qy 7939 ACTTGTCTTCAACCCATCTTTCTTCAAGAACTCTTAGCCCTTATGCGACTACTGCCAA 7998  
Db 2947 TTTGTTTTTCAACCGGTCTACTTCCGAGATTTATGGCCGCTGGAGATCATGACCA 3006  
Qy 7999 CGGGAAGATGTTCTGACAAATTAATTTACATGACCGGTACCGCGGAGAGAGGCGCC 8058  
Db 3007 GGGTCAGCTCGCCCTGAAAACITTTACAGCATCACCGGTTCTGCTGAGAAACGCGGCG 3066  
Qy 8059 TGTCTGACTGGAAGTACTAGCGGATGAGTGTGCGGGAAGCAGCTTGTCCCGATTTC 8118  
Db 3067 AATTACACCGGAGGCTACTGACTGTCTTATCCAGGAAGCGGTCTCATGCCAGATC 3126  
Qy 8119 CTTAATCTGCGATGATGCGAACCGGGCTTCTCCGGCACCCCTGGTGAACCTAGTAA 8178  
Db 3127 TTTAATTTTGCATATCTGAGCGAGGACTGTTGCGGCTTCCATGAGTASCTCAGCGA 3186  
Qy 8179 TGTCTTGAATAATCCAGAAATCATCTCAGCGCCACCCCTGAGTCAATTTGTCTATCAC 8238  
Db 3187 TGTGCTGCAAAAACCCCGCTTATCTGAGGCGCCCTGCGCTGAGCAGTTGTCTATAG 3246  
Qy 8239 AAACAGCACCCGAGTCTGCTCAGTCAAGTCAAGATTTATGCCAACCCAGGGGCTGT 8298  
Db 3247 CAAACCCCATCCCAACATGCGCACACCGTCAGCATCATCCCTTTAAACCC---CATCGG 3303  
Qy 8299 TCCCGGCAATGTTTATGACAGCTGGAAGCGGTGTACAGAGATTCACACAGGC 8358

Db 3304 TACAGACCCGGCGTTTATTAGTACGTGGAGCGCGGTACAGATATGTTGTACACAC 3363  
Qy 8359 ACAGAC 8364  
Db 3364 ATCCAC 3369  
RESULT 7  
US-09-913-514-2  
; Sequence 2, Application US/09913514  
; Patent No. 6653069  
; GENERAL INFORMATION:  
; APPLICANT: GOMI, Yasuyuki  
; APPLICANT: SUNAMACHI, Hiroki  
; APPLICANT: TAKAHASHI, Michiaki  
; APPLICANT: YAMASHI, Koichi  
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vacci  
; FILE REFERENCE: 0216-0454P  
; CURRENT APPLICATION NUMBER: US/09/913,514  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: PCT/JP01/00678  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: JP 2000-62734  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 125157  
; TYPE: DNA  
; ORGANISM: Varicella virus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(125157)  
; OTHER INFORMATION: Attenuated Oka strain  
US-09-913-514-2  
Query Match 1.7%; Score 154; DB 4; Length 125157;  
Best Local Similarity 56.2%; Pred. No. 3.3e-34;  
Matches 289; Conservative 0; Mismatches 225; Indels 0; Gaps 0;  
Qy 585 CAATATTATTGTAATGACGAGCGCGGACCCCTGCTCTTACATACTCACCACCGTCTG 644  
Db 9652 CAAGTTATTGTTATTGATGAGCAGATGCTGAGCGCTCATATCTCACGGCGTGT 96711  
Qy 645 GTTCTTTTACTGTTCTTAAACAGCTGGGTAATAACCCCTTTTACCGCAGGGGCGGT 704  
Db 96712 TTAAGTTGGTGGCTTTTGAATGCTATATCAAAAGCCCTCAGTACATAAAACGGTCAA 96771  
Qy 705 TCCGTGCATAGTATGCTGCTGTTCCGCAACGAGACAAACGGTTCAGTCAACTTACAA 764  
Db 96772 ACCGTCAATAGTATGCTGCTGTTCCGCCACCCAACTGACTGTTAGATCTCATTTCA 96831  
Qy 765 CCACGGAGCGCAAAAGACGAGATATGCTGTCGAGAAACATCTTAAACATTCATAGG 824  
Db 96832 ACATGACATGACGCTTCAACGTAACCTCTAGTGAATAATCTACATATATATCTG 96891  
Qy 825 GAAGAAGTCTGATCTGAGTACATCTGAGAGGAACTGCGGCTGTTTATAAACA 884  
Db 96892 CAATCAACTCTCGCTCAATATATCTAACATCTCAATACTGGCACTTTTATTATTA 96951  
Qy 885 CAAGCGCTGACGATCTGCGCTGCGACACTTCTTAAATAATTTAGAGTATAATCTTCC 944  
Db 96952 CAAACGATGTCAAGAGAGCAATTTTGAATCTTTTAAACCGCTTGTAGTACGGCTACC 97011  
Qy 945 CATTCTGAGGAGTCTGAGTACATGATGATGATGATGATGATGATGATGATGATGAT 1004  
Db 97012 TATTACGAGACATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 97071  
Qy 1005 GGATCTTTTAAATACATTTGGCTGGACCGACTTTTGTTCATACGAGGATTAAGGC 1064  
Db 97072 CAATCTGCTAATCTCCCGGATGAGCGGTCTGTATTCTGCGCATAGGAGGTGAGGC 97131  
Qy 1065 GTAICTAACAAACCTCACACATGCTCTAACGCTA 1098

Db 97132 GTATATGAGTAAGTTACACGGCGCAATTTAAACCTA 97165

## RESULT 8

US-09-661-596A-76  
; Sequence 76, Application US/09661596A  
; Patent No. 6528066  
; GENERAL INFORMATION:  
; APPLICANT: Grose, Charles  
; APPLICANT: Santos, Richard  
; TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE  
; FILE REFERENCE: 140,001,010  
; CURRENT APPLICATION NUMBER: US/09/661,596A  
; CURRENT FILING DATE: 2000-09-14  
; PRIOR APPLICATION NUMBER: US 60/153,779  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 76  
; LENGTH: 124884  
; TYPE: DNA  
; ORGANISM: Varicella zoster  
US-09-661-596A-76

Query Match 1.7%; Score 152.4; DB 4; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1e-33;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;  
QY 585 CAATATTATTGTAATCGACGAGCGCGGAAACCCCTCTCTTACATATCTACACCGCTCGT 644  
Db 96706 CAACGTTATGTTATGATGAGCAGGATTGCTAGGCGTCATATCTCAGCGCGCTGT 96765  
QY 645 GTTCTTTTACTGGTTCCTAAACAGCTGGCTAAATACCCCTCTTTACCGCGAGGGCGGT 704  
Db 96766 TTACTGTTGGTGGCTTTTGAATGCTATATATCAAGCCCTCAGTACATAAAGCTGAAA 96825  
QY 705 TCGTGCATAGTATCGTTCGCTGCGCAACGACGACAAACCGCTTCAGTCAACTTACAA 764  
Db 96826 ACCGTCATAGTATCGTTCGCTGCGCAACGACGACAAACCGTTCAGTCAACTTACAA 96885  
QY 765 CCACGGGACGCAAAAGACGAGATATCGTTCGCGGAGAACATCTCAACATTCATGATAG 824  
Db 96886 ACATGACATGACGGCTTCACACGTAACCTCTAGTGAATAATATACTCAGCTATATACTG 96945  
QY 825 GAAGAGTCTGTAATCTGAGTACACTTGGAGAGAACTGGCGCTGTTTATAACAA 884  
Db 96946 CAATCAAACTCTGGCTCAATATCTAATCATCTCAATCTGAGCAATCTTTTATAATA 97005  
QY 885 CAAGCGCTGCACCGATCTGCAGTTCGACACCTTTCTAAATAATTTAGAGTATAATCTTC 944  
Db 97006 CAAACGATGTCAGAGGACGATTTTGGAAATCTTTTAAACGCTTGAGTACGGCTACC 97065  
QY 945 CATTCCTCAGCAAGTACAGTTCAGTACAGATTTTGTCTTCTTAAAGTAAAGTATTAT 1004  
Db 97066 TATTACCGAGCACATGCGGCTCTGGTGCATACATTTTGTACCTGCATCTCTATATA 97125  
QY 1005 GGATCCTTTAGATATACATGCTGGACCGGACTCTTTTCTCATAGCAGGATAAGGC 1064  
Db 97126 CAATCCTGTAATCTTCCGATGGACGGCTCTGTTCTGTCGATAGGAGTGAAGGC 97185  
QY 1065 GTATCTACAAACCTACACATCTCTAACGCTA 1098  
Db 97186 GTATATGAGTAAGTTACACGGCGCAATTTAAACCTA 97219

## RESULT 9

US-09-913-514-1  
; Sequence 1, Application US/09913514  
; Patent No. 6653069  
; GENERAL INFORMATION:  
; APPLICANT: GOMI, Yasuyuki  
; APPLICANT: SUNAMACHI, Hiroki

; APPLICANT: TAKAHASHI, Michiaki  
; APPLICANT: YAMAMITSU, Koichi  
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vacci  
; FILE REFERENCE: 0216-0454P  
; CURRENT APPLICATION NUMBER: US/09/913,514  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: PCT/JP01/00678  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: JP 2000-62734  
; PRIOR FILING DATE: 2000-01-31  
; NUMBER OF SEQ ID NOS: 42  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 124884  
; TYPE: DNA  
; ORGANISM: Varicella virus  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(124884)  
; OTHER INFORMATION: Dumas Strain  
US-09-913-514-1

Query Match 1.7%; Score 152.4; DB 4; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1e-33;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;  
QY 585 CAATATTATTGTAATCGACGAGCGCGGAAACCCCTCTCTTACATATCTACACCGCTCGT 644  
Db 96706 CAACGTTATGTTATGATGAGCAGGATTGCTAGGCGTCATATCTCAGCGCGCTGT 96765  
QY 645 GTTCTTTTACTGGTTCCTAAACAGCTGGCTAAATACCCCTCTTTACCGCGAGGGCGGT 704  
Db 96766 TTACTGTTGGTGGCTTTTGAATGCTATATCAAGCCCTCAGTACATAAAGCTGCGAAA 96825  
QY 705 TCGTGCATAGTATCGTTCGCTGCGCAACGACGACAAACCGCTTCAGTCAACTTACAA 764  
Db 96826 ACCGTCATAGTATCGTTCGCTGCGCAACGACGACAAACCGTTCAGTCAACTTACAA 96885  
QY 765 CCACGGGACGCAAAAGACGAGATATCGTTCGCGGAGAACATCTCAACATTCATGATAG 824  
Db 96886 ACATGACATGACGGCTTCACACGTAACCTCTAGTGAATAATATACTCAGCTATATACTG 96945  
QY 825 GAAGAGTCTGTAATCTGAGTACACTTGGAGAGAACTGGCGCTGTTTATAACAA 884  
Db 96946 CAATCAAACTCTGGCTCAATATCTAATCATCTCAATCTGAGCAATCTTTTATAATA 97005  
QY 885 CAAGCGCTGCACCGATCTGCAGTTCGACACCTTTCTAAATAATTTAGAGTATAATCTTC 944  
Db 97006 CAAACGATGTCAGAGGACGATTTTGGAAATCTTTTAAACGCTTGAGTACGGCTACC 97065  
QY 945 CATTCCTCAGCAAGTACAGTTCAGTACAGATTTTGTCTTCTTAAAGTAAAGTATTAT 1004  
Db 97066 TATTACCGAGCACATGCGGCTCTGGTGCATACATTTTGTACCTGCATCTCTATATA 97125  
QY 1005 GGATCCTTTAGATATACATGCTGGACCGGACTCTTTTCTCATAGCAGGATAAGGC 1064  
Db 97126 CAATCCTGTAATCTTCCGATGGACGGCTCTGTTCTGTCGATAGGAGTGAAGGC 97185  
QY 1065 GTATCTACAAACCTACACATCTCTAACGCTA 1098  
Db 97186 GTATATGAGTAAGTTACACGGCGCAATTTAAACCTA 97219

## RESULT 10

US-09-651-656-36/c  
; Sequence 36, Application US/09651656  
; Patent No. 6340566  
; GENERAL INFORMATION:  
; APPLICANT: MCCUTHEN-MALONEY, SANDRA  
; APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY  
; TITLE OF INVENTION: DETECTION AND QUANTITATION OF SINGLE NUCLEOTIDE  
; TITLE OF INVENTION: POLYMORPHISMS, DNA SEQUENCE VARIATIONS, DNA MUTATIONS,  
; TITLE OF INVENTION: DNA DAMAGE AND DNA MISMATCHES

FILE REFERENCE: IL-10689  
CURRENT APPLICATION NUMBER: US/09/651,656  
CURRENT FILING DATE: 2000-08-29  
PRIOR APPLICATION NUMBER: 60/192,764  
PRIOR FILING DATE: 2000-03-28  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 36  
LENGTH: 1532  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-09-651-656-36

Query Match  
Best Local Similarity 1.5%; Score 137.6; DB 4; Length 1532;  
Matches 267; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 3638 AGGGAGACGATGCGCGCTTCCAGCAGATGATGCTCTGCTCGGTTTATCAGCGTCGCC 3697  
DB 1104 AGCGGACGATGCGGTGCTTTTCAGTACATGATGGGTTCTATCTATATATCGCCCT 1045  
QY 3698 TTGCTATAGCTTTCTTCCAGCAGATGATGCTCTGCTGAGCTTGGCAGAAAG 3757  
DB 1044 TTCTTTTGGCATGCGATGCCCAACAAAAACACACGCTTCGGGATGCTGTTAATC 985  
QY 3758 CAGCTTATTATGATGCTGGAACCAATCCAGCAGGTTGGCTGTGACCCCGCTTC 3817  
DB 984 AGGCTGATCACTTATCGGTGAAGTTTCCAGCGAGGCTGCGTGGGATCGGCTGA 925  
QY 3818 CCTCTTCCAGGTCAGATGGTGTTPAAAGCAGAACTCCCGTTTGGCCAAAGTCT 3877  
DB 924 CCGCGGTACCGTCAACAGATTTAGTAGCAGAACGCCCTTGACGCGCCAGCTTCA 865  
QY 3878 AAGCACCGGTGAGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAAAATA 3937  
DB 864 AGATTAATGATGAGCGGTGAAGCCCGGTGAGCCGGAATGATTTTCCAGCTTTTATACATA 805  
QY 3938 TTTTGGAGGTGAGGTATATCTAGTCCGAGCAACGCTAAAGCTAGTCCGGTGCT 3997  
DB 804 TTCAATATGACGAGGATGCGAATGCGGAGCAACGAAATGCCAGACCATCGGC 745  
QY 3998 TGACC---GCGATGTAAGATCTCTGGCAGAGGATCAGACTTTATATCTCCGGTTCG 4054  
DB 744 TGTCCCGTCCGATAGGATCTGCGCGAGATCCTGCGCGAGATCACCATTAAAGTACCCTCAACTCT 685  
QY 4055 CAACAATACGACCAACATTAATCTATCTATCTATCGGTGATAAT 4098  
DB 684 GTAAAGCGGAACGCTTAAAGACATCTTTTGTGGGTAGAT 641

RESULT 11  
US-09-650-855-36/c  
Sequence 36, Application US/09650855  
Patent No. 636535  
GENERAL INFORMATION:  
APPLICANT: MCCUTCHEN-MALONEY, SANDRA  
APPLICANT: LAWRENCE LIVERMORE NATIONAL LABORATORY  
TITLE OF INVENTION: CHIMERIC PROTEINS FOR DETECTION AND QUANTIFICATION OF DNA  
TITLE OF INVENTION: MUTATIONS, DNA SEQUENCE VARIATIONS, DNA DAMAGE AND DNA  
TITLE OF INVENTION: MISMATCHES  
FILE REFERENCE: IL-10284  
CURRENT APPLICATION NUMBER: US/09/650,855  
CURRENT FILING DATE: 2000-08-29  
PRIOR APPLICATION NUMBER: 60/192,764  
PRIOR FILING DATE: 2000-03-28  
NUMBER OF SEQ ID NOS: 106  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 36  
LENGTH: 1532  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-09-650-855-36

Query Match  
Best Local Similarity 1.5%; Score 137.6; DB 4; Length 1532;  
Matches 267; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 3638 AGGGAGACGATGCGCGCTTCCAGCAGATGATGCTCTGCTCGGTTTATCAGCGTCGCC 3697  
DB 1104 AGCGGACGATGCGGTGCTTTTCAGTACATGATGGGTTCTATCTATATATCGCCCT 1045  
QY 3698 TTGCTATAGCTTTCTTCCAGCAGATGATGCTCTGCTGAGCTTGGCAGAAAG 3757  
DB 1044 TTCTTTTGGCATGCGATGCCCAACAAAAACACACGCTTCGGGATGCTGTTAATC 985  
QY 3758 CAGCTTATTATGATGCTGGAACCAATCCAGCAGGTTGGCTGTGACCCCGCTTC 3817  
DB 984 AGGCTGATCACTTATCGGTGAAGTTTCCAGCGAGGCTGCGTGGGATCGGCTGA 925  
QY 3818 CCTCTTCCAGGTCAGATGGTGTTPAAAGCAGAACTCCCGTTTGGCCAAAGTCT 3877  
DB 924 CCGCGGTACCGTCAACAGATTTAGTAGCAGAACGCCCTTGACGCGCCAGCTTCA 865  
QY 3878 AAGCACCGGTGAGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAAAATA 3937  
DB 864 AGATTAATGATGAGCGGTGAAGCCCGGTGAGCCGGAATGATTTTCCAGCTTTTATACATA 805  
QY 3938 TTTTGGAGGTGAGGTATATCTAGTCCGAGCAACGCTAAAGCTAGTCCGGTGCT 3997  
DB 804 TTCAATATGACGAGGATGCGAATGCGGAGCAACGAAATGCCAGACCATCGGC 745  
QY 3998 TGACC---GCGATGTAAGATCTCTGGCAGAGGATCAGACTTTATATCTCCGGTTCG 4054  
DB 744 TGTCCCGTCCGATAGGATCTGCGCGAGATCCTGCGCGAGATCACCATTAAAGTACCCTCAACTCT 685  
QY 4055 CAACAATACGACCAACATTAATCTATCTATCTATCGGTGATAAT 4098  
DB 684 GTAAAGCGGAACGCTTAAAGACATCTTTTGTGGGTAGAT 641

RESULT 12  
US-09-489-039A-4593/c  
Sequence 4593, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:  
APPLICANT: Gary Breton et. al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 2709,2004001  
CURRENT APPLICATION NUMBER: US/09/489,039A  
CURRENT FILING DATE: 2000-01-27  
PRIOR APPLICATION NUMBER: US 60/117,747  
PRIOR FILING DATE: 1999-01-29  
NUMBER OF SEQ ID NOS: 14342  
SEQ ID NO 4593  
LENGTH: 723  
TYPE: DNA  
ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-4593

Query Match  
Best Local Similarity 1.5%; Score 134; DB 4; Length 723;  
Matches 267; Conservative 0; Mismatches 200; Indels 3; Gaps 1;

QY 3632 GTAGCCAGGAGACGATGCGCGCTTCAGGACGAGATGCTCTGCTCGGTTTATCAGC 3691  
DB 611 GCAGAGAGCGGTGATGATGCGCGCTTCAGCAGCAGTGGCGCTGAGCATGATATC 552  
QY 3692 GTCGCTTGTCTATAGCTTTCTTCCCAACAGCATAAAAAACGACCTCTGAAGTTGGCA 3751  
DB 551 GCGCTTTCTCTGCGCATGGGAGCCCCACAGCAGAAAAAATACTCCCTCGCAATGCTCA 492  
QY 3752 GAAGGAGCTTATTATGATGCTGGTGAACCAATCCAGCAGGTTGGCTGTGACCCC 3811  
DB 491 TTGATTAAAGCATAACTTATCGGTAAAGGTCTCCAGCCAGACTGCGGATGTC 432



Db 18412 ACTAATAACCAATTTGATGGCAATCCGAAATATCTTGTGTGAGTCTTTATAT 18353  
 QY 3935 ATATTTTTCAGGCTTTGGAGGTATCTGTAGTCCGAGCAACGCTAAAGGCTAGTCCGGTG 3994  
 Db 18352 ATATTTAATAGGAGGGGAATGCTACTTTCAGGTTTTTACTGAAAGCAAGCCGTGC 18293  
 QY 3995 GCTTG---ACCGCGATGTAAAGATCCTGGCCAGGATCAGACTTTTAATATCTCCGGT 4051  
 Db 18292 GCTTGGTTGGTCCATGATAGGATCCTGGACCTAAATATACACTTTTACATCTCTCAAA 18233  
 QY 4052 TCGCAACAATACGACCAACACATAATCTATCTATCGTGGGATTAATC 4099  
 Db 18232 GCAGTATATTGAATGCGTTAAATACATCTTCTTGTGGGGATTAATC 18185

RESULT 15  
 US-09-643-990A-1/c  
 ; Sequence 1, Application US/09643990A  
 ; Patent No. 6528289  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robert D. Fleischmann  
 ; Mark D. Adams  
 ; Owen White  
 ; Hamilton O. Smith  
 ; J. Craig Venter  
 ; TITLE OF INVENTION: The Nucleotide sequence of  
 ; the Haemophilus influenzae Rd Genome, Fragments  
 ; Thereof, and Uses Thereof  
 ; NUMBER OF SEQUENCES: 1  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville,  
 ; STATE: MD  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3 1/2 inch diskette  
 ; COMPUTER: Dell Pentium  
 ; OPERATING SYSTEM: MS DOS v6.22  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/643,990A  
 ; FILING DATE: 23-Aug-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/487,429  
 ; FILING DATE: 1995-06-07  
 ; APPLICATION NUMBER: 08/426,787  
 ; FILING DATE: 1995-04-21  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kenley K. Hoover  
 ; REGISTRATION NUMBER: 40,302  
 ; REFERENCE/DOCKET NUMBER: PB186P1C1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 301-610-5790  
 ; TELEFAX: 310-309-8439  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1830121 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ; US-09-643-990A-1

Query Match 1.3%; Score 117.6; DB 4; Length 1830121;  
 Best Local Similarity 54.7%; Pred. No. 2.8e-22;  
 Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;  
 QY 3635 GCCAGGGGAGCGGATCGCGCTTGGAGCAGAGATGCTCTGTGCTTATCAGGTC 3694

Db 18652 GACACGGGGAGGATGCGGAGCCGTTAAAAACAAGGTGACGAGTGGGTCAATCATTTGC 18593  
 QY 3695 GCCTTGTCTATAGCCTTTCTTCCACACAGATAAAAAACGACCTCTGAAGCTTGGCAGAA 3754  
 Db 18592 CCTTTTTTGTGCGTGACTGCCCCAAAGTAAAAACACAGTTTTTCAAGATGTTCTATG 18533  
 QY 3755 AGCAGCTTATTATGTAGTGTGAAACCAATCCAGCCAAAGGTGGCGGTGACCCCGCC 3814  
 Db 18532 AGTACTGCAATTAATTTATCTGTAAACCTTTCCCAACCTAAATGGCGTGTGAATGTGCC 18473  
 QY 3815 TTCCCTCTTCCACCCGTGAGATGGTGTAAAAAGCAGAACTCCCGCTTTTGGCCCAACAG 3874  
 Db 18472 ATACCTCGTTCCACGGTAAGCAGAGTGTAAAGTAGCAATACCCCTTGTCTGCCCATTTG 18413  
 QY 3875 TCTAAGCACCCGTGAGAAGGAGCGGTGAACCCAGGTACAGTATTTGGCTATCTCTTTAAAA 3934  
 Db 18412 ACTAATAACCAATTTGATGGCAATCCGAAATATCTTGTGTGAGTCTTTATAT 18353  
 QY 3935 ATATTTTGTAGGCTTGGAGGTATATCTGTAGTCCGAGCAACGCTAAAGGCTAGTCCGGTG 3994  
 Db 18352 ATATTTAATAGGAGGGGAATGGGTACTTTCAGGTTTTTACTGAAAGCAAGCCGTGC 18293  
 QY 3995 GCTTG---ACCGCGATGCTAAAGGATCCTGGCCAAAGGATCAGACTTTTAATATCTCCGGT 4051  
 Db 18292 GCTTGGTTGGTCCATGATAGGATCCTGACCTAAATTAACCACTTTTACATCTCTCAAA 18233  
 QY 4052 TCGCAACAATACGACCAACACATAATCTATCTATCGTGGGATTAATC 4099  
 Db 18232 GCAGTATATTGAATGCGTTAAATACATCTTCTTGTGGGGATTAATC 18185

Search completed: April 11, 2004, 08:05:57  
 Job time : 471 secs



```

; SEQ ID NO 2
; LENGTH: 125157
; TYPE: DNA
; ORGANISM: Varicella virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(125157)
; OTHER INFORMATION: Attenuated Oka strain
US-09-913-514-2

Query Match      1.7%; Score 154; DB 10; Length 125157;
Best Local Similarity 56.2%; Pred. No. 4.2e-32;
Matches 289; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 765 CCACGGGACGCAAAAGACGGAGATATCGTCGTGCGAGAACATCCTTAACATTCATGATAGG 824
DB 14344 ACACCAAAAACCTGGATGCTCCGTCGGCAGAGCGGAACGGTCTCACGTACCTCATCG 14285
QY 825 GAAGAAGGTGATCTGAGTACGTACACTTGGAGAGAACTGGCGCCTGTTTATAACAA 884
DB 14284 CAACCGCACCTAGCGAGTACACGGCCCTCTCGCACAGCTGGGCCATTTTCATTAA 14225
QY 885 CAACGGCTGCACCGATCTGCAGTTCGGACACCTTCTAAAAAATTTTAGAGTATAATCTCC 944
DB 14224 CAACGGATGTTGGAGCACAGATTCGGGAACCTCATGAGTCTGAGTACGGCCTTCC 14165
QY 945 CATTCCTGACGAATCATGATGTTAGTACAGATTTGCTGTCCTTAAAGTAAAGTATAT 1004
DB 14164 CATCACCGAGGACACATGATGTTGTGGACCGCTTGTGCTCCCGGAAAGTTACATCA 14105
QY 1005 GGATCTTTTGAATFACATTCGGCTGGACCGCCGACTCTTTTGTACATAGCGAGTAAAGG 1064
DB 14104 CAACCGGGCCAACTTCGGGGGTGGACGGCTGTTCTGTCGCCACAGAGGTACAGCG 14045
QY 1065 GTATCTAACAAACCTACACATGTCTAACGCTAGGGGGCGATACACAGGACACAGAGCT 1124
DB 14044 GTACATGGCCAAAGCTCCACGCTACCTAAAGGTGACTCGCGA---GGGGGAGTTTGT 13988
QY 1125 CTTTACCTGTCCTGCTGTCGGAGGTGTTTGTGAAGCCGTTTGAGAGTACACAGCGGCG 1184
DB 13987 GTTTACCTGTCCTGCTGTCGGAGGTGTTTGTGAAGCCGTTTGAGAGTACACAGCGCT 13928
QY 1185 CGTCAACCTCACCAACCTCACCGTACCGAGTGGGTGACAAAAAATCTTTTAAAGTAA 1244
DB 13927 CACGACGAAACCCAGCTGACCATGGAAGTGGATCACGCCCAACGCCAGTGCATCAC 13868
QY 1245 TAATTTATCGAGTTTGGACGAGGACATGTCTCATGTGTCACGACGATC-----CAC 1298
DB 13867 CAACCTACTCCAGAGTCAGGACGAGGACCGGGGCGACGTCGCTGTGAGGTGCACAGCA 13808
QY 1299 TGAAGTTTCGACCCAGGTTACTTTATCAACCAAGTTTGTAAAAAACGACCGACGTCGTTT 1358
DB 13807 GCACAGTGTGTCGGCCGAGACGACATCAGTACGTCCTCAACAGCCAGGTCGCGGT 13748
QY 1359 AAATGGAACAAACAAATATGATGTGGTTTCAGGGTACGTACTTTAGTTCAAAG 1418
DB 13747 GACCGCGCGCTCCGAAAGATGTTTGGTTTCGACGGGACGTTTCGGACCTTCGAGGC 13688
QY 1419 AATCTAGACGAGCACTCTTCGTCGGAGACGATTCGCAAGACCGTCGGAATACGTATA 1478
DB 13687 TGTGTCGCGACGACAGCTTCGTTGAGAGACCCAGGGGAGACCTCGGTGAGTTCGCTTA 13628
QY 1479 CGGTTTCTTAAACACGCTACTGTATCAACGCGCATGTACTCGTTTCAAC 1524
DB 13627 CCGGTTCTGTCGGGCTCATGTTCCGGCGGCTGATTCACCTTTTAC 13582

```

```

RESULT 2
US-09-913-514-2
; Sequence 2, Application US/09913514
; Publication No. US20030082210A1
; GENERAL INFORMATION:
; APPLICANT: GOMI, Yasuyuki
; APPLICANT: SUNAWACHI, Hiroki
; APPLICANT: TAKAHASHI, Michiaki
; APPLICANT: YAMANISHI, Koichi
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccine
; FILE REFERENCE: 0216-0454P
; CURRENT APPLICATION NUMBER: US/09/913,514
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/JP01/00678
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: JP 2000-62734
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 2
; LENGTH: 125157
; TYPE: DNA
; ORGANISM: Varicella virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(125157)
; OTHER INFORMATION: Attenuated Oka strain
US-09-913-514-2

Query Match      1.7%; Score 154; DB 10; Length 125157;
Best Local Similarity 56.2%; Pred. No. 4.2e-32;
Matches 289; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 585 CAATATATTTCTATCGACGAGCGCGGAACCCCTCTCTTACATATCTCACACCGCTCGT 644
DB 96652 CAAGCTTATTTGTTTATGATGAAGCAGGATTCCTAGGCGGTCATATCTCACGCGCTGT 96711
QY 645 GTTCTTTTACTGTTTCCTTAAACAGCTGGCTAAATACCCCTCTTTACGCCAGGGGGCGGT 704
DB 96712 TTACTGTTGGTGGTTTGTGATATATATCAAGCCCTCAGTACATAAAACGTCGAAA 96771
QY 705 TCCGTGCTATGATGCTGCTGGTTTCCCAACGACGACAAACCGCTTCCAGTCAACTTACAA 764
DB 96772 ACCGCTCATAGTATGCTGCTGGTTTCCCAACGACGACAAACCGCTTCCAGTCAACTTACAA 96831
QY 765 CCACGGGACGCAAAAGACGAGATATCGTCGTGCGAGAACATCCTTAAACATTCATGATAGG 824
DB 96832 ACATGACATGACGCTTCCACAGTAACTCTTAGTGAAATATATCTACGTATATAATCTG 96891
QY 825 GAAGAAGTGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 884
DB 96892 CAATCAAACTCTGCTGCTCAATATCTAACTCACTCACTCACTCACTCACTCACTCACTCA 96951
QY 885 CAAGCGCTGACCGATCTGCTGCGACACCTTCTAAAAAATTTTAGAGTATAATCTTCC 944
DB 96952 CAACGATGTCAGAGGACGATTTTGGAAATCTTTTAAACCGCTTGTAGTACGGCTACC 97011
QY 945 CATTCCTGACGAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1004
DB 97012 TATTACCGAAGCACATGCGCTCTGCTGATACATTTTGTGTACCTGCTCTCTATATATA 97071
QY 1005 GCATCTTTAGATATCATTTGCTGACCGCTCTTTTGTGTCATAGCGAGGTAAAGGC 1064
DB 97072 CAATCTCTGCTATCTCCCGGATGAGCGCTGTATTCGTCGATAGAGGTGAGCGC 97131
QY 1065 GTATCTAAACACCTACACATGCTTAACGCTA 1098
DB 97132 GTATATGATTAAGTTTACACGCGCATTTAAACTA 97165

```

```

RESULT 3
US-09-913-514-1
; Sequence 1, Application US/09913514
; Publication No. US20030082210A1
; GENERAL INFORMATION:
; APPLICANT: GOMI, Yasuyuki
; APPLICANT: SUNAWACHI, Hiroki
; APPLICANT: TAKAHASHI, Michiaki
; APPLICANT: YAMANISHI, Koichi
; TITLE OF INVENTION: Method for Quality Control of an Attenuated Varicella Live Vaccine
; FILE REFERENCE: 0216-0454P
; CURRENT APPLICATION NUMBER: US/09/913,514
; PRIOR FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: PCT/JP01/00678
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: JP 2000-62734
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 124884
; TYPE: DNA

```



ORGANISM: Varicella virus  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(124884)  
OTHER INFORMATION: Dumas Strain  
US-09-913-514-1

Query Match 1.7%; Score 152.4; DB 10; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1.2e-31;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 585 CAATATTATTGTAATCGACGAGCGGGAACCCCTCTTTACATACTCACCACCGTCGT 644  
DB 96706 CAAAGTATTGTTATGATGAAGCAGGATTGCTAGGGCGTCATATCTTCACGGCGGTGT 96765

QY 645 GTTCTTTTACTGTTCTCTAAACAGCTGCCTAAATACCCCTCTTTACGCCAGGGGGCGGT 704  
DB 96766 TTTACTGTTGGTGCTTTTGAATGCTATATATCAAGGCCCTCAGTACATAAAGCGTCAA 96825

QY 705 TCCTGTCATAGTAGTATCGTTCGCCAACGACGACAAAACCGGTTCCAGTCCAATTACAA 764  
DB 96826 ACCGGTCATAGTAGTATCGTTCGCCAACGACGACAAAACCGGTTCCAGTCCAATTACAA 96885

QY 765 CCACGGGACGCAAAAGACGAGATATCGTTCGGGAGAACCTCTTAAAGTCAATCATGATAGG 824  
DB 96886 ACATGACATGCGAGCGTTACACGTAACCTCTAGTGAATAATATACTCACGTTATATCTG 96945

QY 825 GAAGAAGGTCGTATCTGAGTACGTACACTTGGAGAGAACTGGGGCGTGTGTTTTATAACA 884  
DB 96946 CATCAAACTCTCGGTCATATATACTACATCTCACATAAAGTGGCAATCTTTTATAATA 97005

QY 885 CAAAGCGCTGACCGATCTCGAGTTCGGACACCTCTTAAAGTCTTAGAGTATATCTTCC 944  
DB 97006 CAAAGCGATGTCAGAGGACGATTTTGGAAATCTTTTAAAGAACGCTTGAGTACGGGCTACC 97065

QY 945 CATCTCTGACGAAGTCATGAGTTAGTACAGATTTCTCGTCTCTTAAAGTAAAGATTAT 1004  
DB 97066 TATTACGAGACACATGCGGCTCTGTCGATACATTTCTGACCTGCATCTATATTA 97125

QY 1005 GGATCTCTTAGAATACATTTGGCTGGACCGGACTCTTTTGTACATACGAGGTAAAGGC 1064  
DB 97126 CAATCTCTGCTAAATCTTTCCCGGATGGACGCGTCTGTTTCTGCGATAAGGAGTGC 97185

QY 1065 GTATCTAAACAAACCTACACATGCTTAACGCTA 1098  
DB 97186 GTATGCTAGTACGCGCATTTAAAACTA 97219

RESULT 5  
US-10-369-493-47194/c  
Sequence 47194, Application US/10369493  
Publication No. US20030233675A1  
GENERAL INFORMATION:  
APPLICANT: Cao, Yongwei  
APPLICANT: Hinkle, Gregory J.  
APPLICANT: Slater, Steven C.  
APPLICANT: Goldman, Barry S.  
APPLICANT: Chen, Xianfeng  
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
FILE REFERENCE: 38-10(52052)B  
CURRENT FILING DATE: 2003-02-28  
PRIOR APPLICATION NUMBER: US 60/360,039  
PRIOR FILING DATE: 2002-02-21  
NUMBER OF SEQ ID NOS: 47374  
SEQ ID NO 47194  
LENGTH: 690  
TYPE: DNA  
ORGANISM: Escherichia coli  
US-10-369-493-47194

Query Match 1.5%; Score 137.6; DB 15; Length 690;  
Best Local Similarity 57.5%; Pred. No. 4.3e-29;  
Matches 267; Conservative 0; Mismatches 194; Indels 3; Gaps 1;

QY 3638 AGGGAGACCGATGGCGCGCTTGGAGGAGAGATGCTCTCTCCCTTTATAGCGTCCGCC 3697  
DB 572 AGCGGACGAGTGGCTGCTTTCAGTACATGATGCGTGTGTTATCTATATAATCGCCCT 513

QY 3698 TTGCTATAGCTTTCTTCCGACAGCATAAAAACGACCTCTGAAGCTTGCACAGAAAG 3757

ORGANISM: Varicella virus  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)..(124884)  
OTHER INFORMATION: Dumas Strain  
US-09-913-514-1

Query Match 1.7%; Score 152.4; DB 10; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1.2e-31;  
Matches 288; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

QY 585 CAATATTATTGTAATCGACGAGCGGGAACCCCTCTTTACATACTCACCACCGTCGT 644  
DB 96706 CAAAGTATTGTTATGATGAAGCAGGATTGCTAGGGCGTCATATCTTCACGGCGGTGT 96765

QY 645 GTTCTTTTACTGTTCTCTAAACAGCTGCCTAAATACCCCTCTTTACGCCAGGGGGCGGT 704  
DB 96766 TTTACTGTTGGTGCTTTTGAATGCTATATATCAAGGCCCTCAGTACATAAAGCGTCAA 96825

QY 705 TCCTGTCATAGTAGTATCGTTCGCCAACGACGACAAAACCGGTTCCAGTCCAATTACAA 764  
DB 96826 ACCGGTCATAGTAGTATCGTTCGCCAACGACGACAAAACCGGTTCCAGTCCAATTACAA 96885

QY 765 CCACGGGACGCAAAAGACGAGATATCGTTCGGGAGAACCTCTTAAAGTCAATCATGATAGG 824  
DB 96886 ACATGACATGCGAGCGTTACACGTAACCTCTAGTGAATAATATACTCACGTTATATCTG 96945

QY 825 GAAGAAGGTCGTATCTGAGTACGTACACTTGGAGAGAACCTGGGGCGTGTGTTTTATAACA 884  
DB 96946 CATCAAACTCTCGGTCATATATACTACATCTCACATAAAGTGGCAATCTTTTATAATA 97005

QY 885 CAAAGCGCTGACCGATCTCGAGTTCGGACACCTTCTTAAAGTCTTAGAGTATATCTTCC 944  
DB 97006 CAAAGCGATGTCAGAGGACGATTTTGGAAATCTTTTAAAGAACGCTTGAGTACGGGCTACC 97065

QY 945 CATCTCTGACGAAGTCATGAGTTAGTACAGATTTCTCGTCTCTTAAAGTAAAGATTAT 1004  
DB 97066 TATTACGAGACACATGCGGCTCTGTCGATACATTTCTGACCTGCATCTATATTA 97125

QY 1005 GGATCTCTTAGAATACATTTGGCTGGACCGGACTCTTTTGTACATACGAGGTAAAGGC 1064  
DB 97126 CAATCTCTGCTAAATCTTTCCCGGATGGACGCGTCTGTTTCTGCGATAAGGAGTGC 97185

QY 1065 GTATCTAAACAAACCTACACATGCTTAACGCTA 1098  
DB 97186 GTATGCTAGTACGCGCATTTAAAACTA 97219

RESULT 4  
US-10-288-823-76  
Sequence 76, Application US/10288823  
Publication No. US20030166168A1  
GENERAL INFORMATION:  
APPLICANT: Grose, Charles  
APPLICANT: Santos, Richard  
APPLICANT: Viant  
TITLE OF INVENTION: VARIANT VARICELLA-ZOSTER VIRUSES AND METHODS OF USE  
FILE REFERENCE: 140-00110102  
CURRENT APPLICATION NUMBER: US/10/288,823  
CURRENT FILING DATE: 2002-11-06  
PRIOR APPLICATION NUMBER: US 09/661,596  
PRIOR FILING DATE: 2000-09-14  
PRIOR APPLICATION NUMBER: US 60/153,779  
PRIOR FILING DATE: 1999-09-14  
NUMBER OF SEQ ID NOS: 80  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 76  
LENGTH: 124884  
TYPE: DNA  
ORGANISM: Varicella zoster  
US-10-288-823-76

Query Match 1.7%; Score 152.4; DB 14; Length 124884;  
Best Local Similarity 56.0%; Pred. No. 1.2e-31;

Db 512 TTCTTTGGCGATCGATCCCAACAAAAACACACGCGCTTCGGGATGCTGGTTAATC 453  
Qy 3758 CAGCTTATTATGATGTTGTAACCAATCCAGCCAGGTTGGCGTGTACCCCGCTTC 3817  
Db 452 AGGCTGATCACTTTATCGGTGAAGTTTCCAGCCAGGCTGGCGTGGGAATGCGCGCTGA 393  
Qy 3818 CCTCTTTCCACCGTCAGAAATGGTGTAAAGCAGAACTCCCGGTTTTCCCAACAGTCT 3877  
Db 392 CCTGGCGGTACCGTCAACACAGTATTGAGTAGCAGAACGCGCTGACGCGCCAGCTTTC 333  
Qy 3878 AAGCACCGCTGAGAACGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAATA 3937  
Db 332 AGATAACCATGATTAGGCGCGGTGAAGCCCGGAATAGTATTTCCAGCTCTTTATACATA 273  
Qy 3938 TTTTTCAGGCTTCGAGGTATACGTAGTCCGAGCAACGCTAAAGGCTAGTCCGCTGGCT 3997  
Db 272 TTCAATATGACGGAGGAATGGCAATGCGCGGACGAAATGCCAGACCATCGCC 213  
Qy 3998 TGACC---GGATGGTAAGATCTGCGCCAAAGATCACGACTTTATATATCTCCCGGTTCG 4054  
Db 212 TGTCCCGGTCGTTGATAAGGATCTGCGCCGAGAATCACCACTTTAAACGTCAOCCAACTCT 153  
Qy 4055 CAACAAATACGACCAACCAATAATCTATCTATCTATCGGTGGATAAAT 4098  
Db 152 GTAAAGCGGAACGGCTTAAAGACATCTTTTGTGGGTAGAT 109

RESULT 6  
US-10-260-877-13/c  
; Sequence 13, Application US/10260877  
; Publication No. US20030021813A1  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Chovan, Linda E.  
; APPLICANT: Hessler, Paul E.  
; APPLICANT: Reich, Karl A.  
; TITLE OF INVENTION: ESSENTIAL BACTERIA GENES AND GENOME  
; TITLE OF INVENTION: SCANNING IN HAEMOPHILUS INFLUENZAE FOR THE IDENTIFICATION OF  
; TITLE OF INVENTION: 'ESSENTIAL GENES'  
; FILE REFERENCE: 6565.US.P1  
; CURRENT APPLICATION NUMBER: US/10/260,877  
; CURRENT FILING DATE: 2002-09-30  
; PRIOR APPLICATION NUMBER: US/09/649,145  
; PRIOR FILING DATE: 2000-08-25  
; NUMBER OF SEQ ID NOS: 137  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 657  
; TYPE: DNA  
; ORGANISM: H. influenzae  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(657)  
; OTHER INFORMATION: HL-0018  
US-10-260-877-13

Query Match 1.3%; Score 117.6; DB 14; Length 657;  
Best Local Similarity 54.7%; Pred. No. 3.6e-23;  
Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;

Qy 3635 GCGAGGGAGACGGAATGCGGCGCTTGAGGACGAGATGCTCTGTCTCGTTTATCAGCGTC 3694  
Db 566 GACAACGGGAAGGATGCGGAGCGCTTAAACAAAGGTGACGAGTGGCGGTCAATCATTTGC 507  
Qy 3695 GCCTTGCTATAGCTTTCTTCCCAACAGATATAAAACGCACTCTGAAGCTTGGCAGAA 3754  
Db 506 CCTTTTCTTTTGTGGTGACTGCCCCCAAGTAATAAACACCACTTTTTCACGATGTTCAATTG 447  
Qy 3755 AGGACGCTTATTATGATGCTGGTGAACCAATCCAGCCAGGTTGGCGTGTGACCCGCC 3814  
Db 446 AGTACTGCAATAAATTTATCTGTAAACCTTTCCCAACCTAAATTTGGCGTGTGAATGTC 387  
Qy 3815 TTCCCTCTTTCCACCGTCAGAAATGCTGTTTTAAAGACGAACTCCCGCTTTTGCCCAACAG 3874

Db 386 ATACCTCGTTCACCGTAAGCACAGTGTAAAGTACCAATACCCCTTGTCTGCCCATTTG 327  
Qy 3875 TCTAAGCACCGCTGAGAGGAGCGGTGAACCCAGGTACAGTATGCTCTTTAAAA 3934  
Db 326 ACTAATAACCATTTGATGGCAATTTGAATCCCGAATACTTGTGTGAGTCTTTATAT 267  
Qy 3935 ATATTTTGGAGTTGAGGTATATCTAGTCCGAGCAACGCTAAAGGCTAGTCCGCTG 3994  
Db 266 ATATTAATAGGAAGGGAATGCTACTTACGTTTACTGAAAAAGCCAGCCGTC 207  
Qy 3995 GCTTG---ACGCGCATGTAAGGATCTGGCCAAAGGATCAGACTTTTAATATCTCCGGT 4051  
Db 206 GCTTGGTTGGTCCATGATAAGGATCTGACCTAAATATACCCTTTTACATCTCAAA 147  
Qy 4052 TCGAACAATAACGACCAACCAATAATTTCTATCTATCGGTGGATAAATC 4099  
Db 146 GCGATATATTGTAATCGGTAAATACATCTTCTTGTGGGGATAAATC 99

RESULT 7  
US-10-329-960-1/c  
; Sequence 1, Application US/10329960  
; Publication No. US20030099277A1  
; GENERAL INFORMATION:  
; APPLICANT: Fleischmann et al.  
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fra  
; TITLE OF INVENTION: Thereof, and Uses Thereof  
; FILE REFERENCE: PBI86P1  
; CURRENT APPLICATION NUMBER: US/10/329,960  
; CURRENT FILING DATE: 2003-01-02  
; PRIOR APPLICATION NUMBER: US 09/643,990  
; PRIOR FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: US 08/487,429  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: US 08/426,787  
; PRIOR FILING DATE: 1995-04-21  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1830121  
; TYPE: DNA  
; ORGANISM: Haemophilus influenzae  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (4747)..(4747)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (9921)..(9921)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (10150)..(10150)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (29298)..(29298)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36543)..(36543)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36551)..(36551)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (36636)..(36636)  
; OTHER INFORMATION: n equals a, t, g or c  
; FEATURE:  
; NAME/KEY: misc feature

```
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c

LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (133360)..(133360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
```

```

RESULT 8
US-10-329-670-1/c
; Sequence 1, Application US/10329670
; Publication No. US20040018503A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186PI
; CURRENT APPLICATION NUMBER: US/10/329,670
; CURRENT FILING DATE: 2002-12-24
; PRIOR APPLICATION NUMBER: US 09/643,990
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: US 08/487,429
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, g or c
; FEATURE:

```

LOCATION: (65313)..(65313)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (80024)..(80024)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (100091)..(100091)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (102696)..(102696)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (105121)..(105121)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (107248)..(107248)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (117136)..(117136)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (119750)..(119750)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (119924)..(119924)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (120038)..(120038)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (121344)..(121344)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (122167)..(122167)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (122336)..(122336)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (131340)..(131340)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (131360)..(131360)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (139910)..(139910)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (140398)..(140398)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (142750)..(142750)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (145058)..(145058)

OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (145171)..(145171)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (145942)..(145942)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (147197)..(147197)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (150841)..(150841)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (152500)..(152500)  
OTHER INFORMATION: n equals a, t, g or c  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (152530)..(152530)

Query Match 1.3%; Score 117.6; DB 15; Length 1830121;  
Best Local Similarity 54.7%; Pred. No. 2.4e-20;  
Matches 256; Conservative 0; Mismatches 209; Indels 3; Gaps 1;  
QY 3635 GCACGGGAGACGGATGGCGCCCTTGAGGACGAGATGTCTGTCCCTTTATCAGCTC 3694  
Db 18652 GACACGGGGAAGGATGGCGGACCGTTAAACAAAGTGACGAGTCGGTCAATCTTGC 18593  
QY 3695 GCCTTGTCTATAGCCTTTCTCCCAACAGCATAAACAGCCACCTCTGAAGCTTGGCAGAA 3754  
Db 18592 CCTTTTTTTGTGGTGACTGCCCCCAAGTAAACACACAGTTTTTACGATGTTCAATTG 18533  
QY 3755 AGCAGCTTATATGTAGCTGTGAACCAATCCAGCAAGGTGGCGGTGTGACCCGCC 3814  
Db 18532 AGTACTGCAATAACTTTATCTGTAAACCTTTCCCAACCTAAATGGCGGTGTAATGTGCC 18473  
QY 3815 TTCCCTCTTTCCACCGTCAGATGGTGTAAACAGCAAGTCTCCCGCTTTTGGCCCAACAG 3874  
Db 18472 ATACCTCGTTCCACGGTAAGCACAGTGTAAAGTASCAATACCCCTTGTCTGCCCATTTG 18413  
QY 3875 TCTAGCACCCGTGAGAGGACGGTGAACCAAGTACAGTATGGCTATCTCTTTAAA 3934  
Db 18412 ACTAAATAACCATTTGATGGCATTTTGAATCCGGAATATCTGTGTAGTCTTTTATAT 18353  
QY 3935 ATATTTTGTAGGCTTGGAGGTATCTGTAGTCCGAGCAAGCTAAAGGCTAGTCCGGTG 3994  
Db 18352 ATATTTATAGGAAGGGGATGGCTACTTCAGSTTTTACTGAAAGCAAGCCGCTG 18293  
QY 3995 GCTTG---ACCGGATGTAGGATCCTGGCCAGGATCAGACTTAAATATCTCCCGT 4051  
Db 18292 GCTTGTTTGGTCCATGATAGGATCCTGACCTAAATACCACTTTTACATCTCTCAAAA 18233  
QY 4052 TCGCAACATACGACCAACCAATATCTATCTATCGGTGGATAAATC 4099  
Db 18232 GCAGTATATTGATGCGTTAATACATCTTCTTGTGGGGGATAAATC 18185

## RESULT 9

US-10-369-493-44794/c  
; Sequence 44794, Application US/10369493  
; Publication No. US2003023675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

;; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
;; FILE REFERENCE: 38-10(52052)B  
;; CURRENT APPLICATION NUMBER: US/10/369,493  
;; PRIOR FILING DATE: 2003-02-28  
;; PRIOR APPLICATION NUMBER: US 60/360,039  
;; PRIOR FILING DATE: 2002-02-21  
;; NUMBER OF SEQ ID NOS: 47374  
;; SEQ ID NO 44794  
;; LENGTH: 675  
;; TYPE: DNA  
;; ORGANISM: Xenorhabdus nematophilus  
US-10-369-493-44794

Query Match 1.2%; Score 110.6; DB 15; Length 675;  
Best Local Similarity 53.5%; Pred. No. 4.3e-21;  
Matches 255; Conservative 0; Mismatches 219; Indels 3; Gaps 1;  
QY 3635 GCCAGGAGAGGATGCGCGCTTGGAGCAGAGATGCTCTGTCGGTTATCAGCGTC 3694  
DB 575 GACAAAGAGAGGAGTGGTGGCTTTTAGAACATAGTGTGTTGGGTATTGATGAAGTGT 516  
QY 3695 GCCTTGCTATAGCCTTTCTTCCCCACAGCATAAACCGCACCTCTGAAGCTTGGCAGAA 3754  
DB 515 CTTTTTTCTGGCATGGGAACCCCAAGTAGGAATAATGACGCGTGACATGTTTCATG 456  
QY 3755 AGGACGCTATATGACTGGTGAACCAATCCAGCAGAGTGGGGTGTACCCCGCC 3814  
DB 455 ATTGCCGCAATGACTTTGTGCTAGTGAAGTTTCCAGCCTAAATTCGCATGTGAATGCCA 396  
QY 3815 TTCCCTCTTCCACCGTCAGAAATGGTGTAAAGAGCAGAACTCCCGGTTTGGCCCAACAG 3874  
DB 395 TTGCCACGTTCAACGGTCAGCACCGTATTGAGCAGCAATACGCCCTGCTTTGCCAACG 336  
QY 3875 TCTAAGACCGGTGAGAAAGGCGGTGAACCAATCCAGCAGTACATATGCTATCTTTAAAA 3934  
DB 335 ACCAAGACGCGATGATTAGGCGCGCTTAAACCCAGTAAATCCGATTCACGCTCTTTATAC 276  
QY 3935 ATATTTTGGGCTTGGAGGTATATCTAGTCCGAGCAACGCTAAAGGCTAGTCCGCGTG 3994  
DB 275 ATGTTGACAGAGAGGTGTGTCAGGAATACCGAGCTGTACTGAGAAAGCGACCGTGT 216  
QY 3995 GCTTGA---CGCGATGTAGAGTCTCTGGCAGAGATCAGACTTAAATCTCCGCT 4051  
DB 215 GCCTGATTGGCGCGTGGTAGGGTCTTGGCCCAAAATCAGACTCTTATCATCGGACGT 156  
QY 4052 TCGCAACAATACGACCAACCATATTTCTATCTATCGGTGGATAAAATCAGGGTTAGC 4108  
DB 155 TGGTATAGCGGAATGCAATTAAGAAACATCTCTGCTGTAGCGGATAAATGCTTTCCG 99

RESULT 10  
US-09-758-017A-1/c  
; Sequence 1, Application US/09758017A  
; Patent No. US20020155573A1  
; GENERAL INFORMATION:  
; APPLICANT: Lanes, Olav  
; APPLICANT: Willasen, Nils Peder  
; APPLICANT: Guddal, Per Henrik  
; APPLICANT: Gjellesvik, Dag Rune  
; TITLE OF INVENTION: Cod uracil-DNA glycosylase, gene coding therefore,  
; TITLE OF INVENTION: recombinant DNA containing said gene or operative parts  
; TITLE OF INVENTION: thereof, a method for preparing said protein and the  
; TITLE OF INVENTION: use of said protein or said operative pa  
; FILE REFERENCE: U013209-3  
; CURRENT APPLICATION NUMBER: US/09/758,017A  
; CURRENT FILING DATE: 2001-01-10  
; PRIOR APPLICATION NUMBER: 2000 5428  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: 2000 0163  
; PRIOR FILING DATE: 2000-01-12  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1

;; LENGTH: 1283  
;; TYPE: DNA  
;; ORGANISM: Gadus morhua  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (18)..(920)  
US-09-758-017A-1

Query Match 1.2%; Score 107.4; DB 9; Length 1283;  
Best Local Similarity 51.4%; Pred. No. 6.6e-20;  
Matches 274; Conservative 0; Mismatches 256; Indels 3; Gaps 1;  
QY 3632 GTAGCCAGGGAGACGATGGCGCTTGGAGCAGAGATGTCTCTGTCGGTTTATCAGC 3691  
DB 829 GCAGACAAAGAGATGATGAACAGCTTGCACACATGCTGAGCTTCTCTGTCGATGCTC 770  
QY 3692 GTGCGCTTGTCTATAGACCTTTCTTCCCGCAGCAGTAAACGACCTCTGAGACTTGGCA 3751  
DB 769 GCTCCCTTCTTATGCGGCTATAGCCCCCAACAGGAAACGACTCTTCCCGTTGACG 710  
QY 3752 GAAAGGAGCAGTTATATGAGCTGTGTAACCAATCCAGCCAAAGGTGGCGTGTGACCCC 3811  
DB 709 CTCAGCCACTTGATCAGCGTCGGTGAAGGTCTCCAGCCTCTGCTCTTGTGGAGTTG 650  
QY 3812 GCCTTCCCTTTTCCACCGTCAGAAATGGTGTAAAGCAGAACTCCCGGTTTGGCCAA 3871  
DB 649 GCCTGATGGCGCCAGCGTCAAGCAGCGGTAAAGCAGCAGCACCCTTGTGTTGCCCCAT 590  
QY 3872 CAGTCTAAGCACCCGCTGAGAAAGGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTA 3931  
DB 589 CCGCTTAGATCTCCATGTCAGGATGCTTGAAGCCATCAATGTCGTACACAATCTTTG 530  
QY 3932 AATAATTTTTGAGGCTTGGAGTACTATGTTAGTCCGAGCAACGCTAAGGCTAGTCCG 3991  
DB 529 TATATGTTACAGAGACTGGGGGAGGGGAACTGGCTTTTGCACACTGAAACAGAGTCCA 470  
QY 3992 GTGCTTGACCGGA---TGTTAAGGATCTCTGCGCAAGGATCAGCACTTTAATATCTCTC 4048  
DB 469 TGTGCTTGTGGACCGTGTAAAGGTCTCTGCGCTTGAATCACTATTTCACATCTTGA 410  
QY 4049 GGTTCGCAACAATACGACCAACACATAATCTCTATCTGTTGTAATCAATCAGGTTAGC 4108  
DB 409 ATGTACACATCTCTGCGAACTGTACACTTGATCAGCCGTTGGGTAGACGTTGACGG 350  
QY 4109 TTGTCATATCCATAACGCGCTTTCAGCAGAGCGCGAGCTTCTGTTTAAAAA 4161  
DB 349 CTCCTCTCATCAGCTACAAAGGACATCAATGTTTGAAGTATGGCTTTTCAAA 297

RESULT 11  
US-10-369-493-24211/c  
; Sequence 24211, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 24211  
; LENGTH: 744  
; TYPE: DNA  
; ORGANISM: Deinococcus radiodurans  
US-10-369-493-24211





QY 3995 GCTTG-----ACCGCATGTAAGATCCTGCCAGGATCAGACTTTA 4039  
Db 149 GCTTGGCGGTGCCACGGCTTCGCGTGATACGGATCCTGCCAGGATCACCACCTTC 90  
QY 4040 ATATCTCCGG 4050  
Db 89 ACTTCATCGG 79

## RESULT 14

US-10-369-493-31014/c  
; Sequence 31014, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 31014  
; LENGTH: 987  
; TYPE: DNA  
; ORGANISM: Burkholderia cepacia  
US-10-369-493-31014

Query Match 1.2%; Score 104.6; DB 15; Length 987;  
Best Local Similarity 56.6%; Pred. No. 3.6e-19;  
Matches 244; Conservative 0; Mismatches 169; Indels 18; Gaps 2;  
QY 3635 GCACGGGAGACGGATGCGCGCTTGAGGACAGATGCTCTGTCGGTTATCAGCGTC 3694  
Db 902 GACAGGGGACGGATGCGGTGTCTCCAGACGCAATGGACTTGCCTCCGCGGAGCGCG 843  
QY 3695 GCCTTGCTATAGCTTTCTTCCACACAGCATAAACACACCTCTGAAGCTTGGCAGAA 3754  
Db 842 CGCTTGCTTGCATGCGGCGCCACAGCATGAACACAGGCGCGTCATGGCGCATCGG 783  
QY 3755 AGGAGCTTATATGATGCTGTGAACCAATCCAGCCAGGTTGCGGTGACCCCGCC 3814  
Db 782 AGTTCGTGATCAGCGTGTGCGTGCAATTTTCCAGCGCGGTTTTCGCTGCGTGGCGCC 723  
QY 3815 TTCCCTCTTCCACCGTCAGATGGTGTAAAGCAGAACTCCCGTTTTCGCCAACAG 3874  
Db 722 GAGTCTGCTCCACCGTCACACCGTGTTCAGCAGCAGCAGCGCTTCCGCGCAGGTG 663  
QY 3875 TCTAAGCACCGGTGAAGAGCGGTGAACCCAGGTACAGTATGCGTATCTCTTTAAA 3934  
Db 662 TCGAGACAGCGGTGAACCGCGCTCTCGTGCCCA--AGGCTTGGCGCGATCTCTTTGAA 606  
QY 3935 ATATTTTGGGCTTGAGGTATCTAGTCCGAGCAACGCTAAAGCTAGTCGCGTG 3994  
Db 605 ATGTTGCGGACGACGCGCGGTGCGCATCTTCGGCGCTACCGAAACCGCGCGCGTGC 546  
QY 3995 GCTTG-----ACCGCATGTAAGGATCTCTGGCCAAAGGATCAGACTTTA 4039  
Db 545 GCTTGGCGGTGCCACGGTCTTCGCGTGATACGGATCCTGGCGGAGGATCACCACCTTC 486  
QY 4040 ATATCTCCGG 4050  
Db 485 ACTTCATCGG 475

## RESULT 15

US-10-369-493-32380/c  
; Sequence 32380, Application US/10369493

; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 32380  
; LENGTH: 675  
; TYPE: DNA  
; ORGANISM: Ralstonia metallidurans  
US-10-369-493-32380

Query Match 1.1%; Score 101; DB 15; Length 675;  
Best Local Similarity 54.8%; Pred. No. 3.1e-18;  
Matches 200; Conservative 0; Mismatches 165; Indels 0; Gaps 0;  
QY 3635 GCACGGGAGACGGATGCGCGCTTGAGGACAGATGCTCTGTCGGTTATCAGCGTC 3694  
Db 575 GACAGCGGACGGATGCGCGCTTCCAGACGCAATGCGTCTGGCCTTCGAGCAAGCG 516  
QY 3695 GCCTTGCTATAGCTTTCTTCCACACAGCATAAACACGCACTCTGAAGCTTGGCAGAA 3754  
Db 515 CGCTTGCGCTGCGGTGCTGCCAGAGCAGGAGGACAGGATTCGGCCGCGATGCCGCC 456  
QY 3755 AGGAGCTTATATGATGCTGTGAACCAATCCAGCAAGGTTGGCGTGTGACCCCGCC 3814  
Db 455 AGCGGTGATCAGCAGTGGTTCACAGCTTCCAGCGCGCTTCGCATGACTGGCGGCC 396  
QY 3815 TTCCCTCTTCCACCGTCAGATGGTGTAAAGCAGAACTCCCGTTTTCGCCAACAG 3874  
Db 395 TGTCCCTGTGACCGTCACTACCGTGTTCAGCAGCAGCAGCGCTGACGTGCCAGCCT 336  
QY 3875 TCTAAGCACCGGTGAAGAGCGGTGAACCCAGGTACAGTATGGCTATCTCTTTAAA 3934  
Db 335 TCCAGATTGCCGACGACGCGTCCGCGCCACGCGCTATTCCGCGCGGATCTCTCTGAAG 276  
QY 3935 ATATTTTGGGCTTGAGGTATCTAGTCCGAGCAACGCTAAAGGCTAGTCCCGTG 3994  
Db 275 ATGTTGCGAGGCTTGGTGGCACCTTGACGCCCTTCGGGTACGAGAAAGCGCCGCTGT 216  
QY 3995 GCTTG 3999  
Db 215 GCCTG 211

Search completed: April 11, 2004, 13:37:24  
Job time : 2015 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 10, 2004, 20:47:25 ; Search time 13557 Seconds  
(without alignments)  
19985.215 Million cell updates/sec

Title: US-09-831-000-1\_COPY\_60966\_70038

Perfect score: 9073

Sequence: 1 atggagagtcctcgatg.....ggggccacgcatgcgttga 9073

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_estl:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_est3:\*

12: gb\_est4:\*

13: gb\_est5:\*

14: gb\_est6:\*

15: em\_estom:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pla:\*

20: em\_gss\_arc:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 241.6 | 2.7         | 717    | 29 | AG101978 Pan trogl |
| 2          | 233   | 2.6         | 678    | 29 | AG104414 Pan trogl |
| 3          | 213   | 2.3         | 688    | 29 | AG126121 Pan trogl |
| 4          | 194.4 | 2.1         | 721    | 29 | AG051826 Pan trogl |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

|    |       |     |      |    |          |                    |
|----|-------|-----|------|----|----------|--------------------|
| 5  | 135   | 1.5 | 685  | 29 | AG124918 | AG124918 Pan trogl |
| 6  | 111   | 1.2 | 656  | 29 | AG051825 | AG051825 Pan trogl |
| 7  | 105.8 | 1.2 | 696  | 9  | AJ395572 | AJ395572 AJ395572  |
| 8  | 103.6 | 1.1 | 911  | 13 | EU444447 | EU444447 603213237 |
| 9  | 103.6 | 1.1 | 987  | 13 | EX706705 | EX706705 BX706705  |
| 10 | 103.2 | 1.1 | 871  | 13 | EX719507 | EX719507 BX719507  |
| 11 | 103   | 1.1 | 645  | 9  | AL894178 | AL894178 AL894178  |
| 12 | 102.4 | 1.1 | 699  | 14 | CF521628 | CF521628 AGENCOURT |
| 13 | 101   | 1.1 | 806  | 9  | AL669156 | AL669156 AL669156  |
| 14 | 100.8 | 1.1 | 574  | 12 | BI940552 | BI940552 db71f06.Y |
| 15 | 100.8 | 1.1 | 687  | 14 | CD321504 | CD321504 StrPUS38  |
| 16 | 100.8 | 1.1 | 935  | 14 | CA983550 | CA983550 AGENCOURT |
| 17 | 100.2 | 1.1 | 963  | 13 | BU442636 | BU442636 603210192 |
| 18 | 100   | 1.1 | 823  | 14 | CB938158 | CB938158 IPCGX13   |
| 19 | 96.6  | 1.1 | 658  | 12 | BJ527177 | BJ527177 BJ527177  |
| 20 | 94.8  | 1.0 | 909  | 13 | EX754661 | EX754661 BX754661  |
| 21 | 93.8  | 1.0 | 822  | 13 | EX744303 | EX744303 BX744303  |
| 22 | 93.4  | 1.0 | 840  | 14 | CD643361 | CD643361 AGENCOURT |
| 23 | 92.2  | 1.0 | 717  | 14 | CD764451 | CD764451 GBSZLBI01 |
| 24 | 92    | 1.0 | 612  | 12 | BJ064298 | BJ064298 BU064298  |
| 25 | 91.2  | 1.0 | 609  | 9  | AU126319 | AU126319 AU126319  |
| 26 | 91    | 1.0 | 540  | 14 | CF847247 | CF847247 pshB042xg |
| 27 | 90    | 1.0 | 822  | 12 | BI823926 | BI823926 603039103 |
| 28 | 90    | 1.0 | 871  | 13 | BQ228775 | BQ228775 AGENCOURT |
| 29 | 90    | 1.0 | 892  | 13 | EU156625 | EU156625 AGENCOURT |
| 30 | 90    | 1.0 | 1052 | 12 | EM541301 | EM541301 AGENCOURT |
| 31 | 90    | 1.0 | 1175 | 12 | EM471137 | EM471137 AGENCOURT |
| 32 | 89.8  | 1.0 | 773  | 14 | CF947286 | CF947286 UI-D-GC1- |
| 33 | 89.4  | 1.0 | 962  | 14 | CA973075 | CA973075 AGENCOURT |
| 34 | 88.8  | 1.0 | 622  | 12 | BI066406 | BI066406 pglfn.pk0 |
| 35 | 88.4  | 1.0 | 548  | 9  | AA290918 | AA290918 zs4b06.r  |
| 36 | 88.2  | 1.0 | 871  | 13 | EX423498 | EX423498 BX423498  |
| 37 | 88    | 1.0 | 1201 | 13 | EX342506 | EX342506 BX342506  |
| 38 | 88    | 1.0 | 1201 | 13 | EX379137 | EX379137 BX379137  |
| 39 | 87.4  | 1.0 | 736  | 12 | BJ584335 | BJ584335 BX584335  |
| 40 | 87    | 1.0 | 606  | 13 | BQ826724 | BQ826724 Gd46b09.X |
| 41 | 85.6  | 0.9 | 786  | 13 | EX887665 | EX887665 BX887665  |
| 42 | 84.8  | 0.9 | 669  | 9  | AU039323 | AU039323 AU039323  |
| 43 | 84.8  | 0.9 | 757  | 12 | BJ373039 | BJ373039 BJ373039  |
| 44 | 84.2  | 0.9 | 913  | 13 | BQ882986 | BQ882986 AGENCOURT |
| 45 | 84    | 0.9 | 937  | 14 | CA983409 | CA983409 AGENCOURT |

## ALIGNMENTS

### RESULT 1

AG101978

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

AG101978 717 bp DNA linear GSS 03-NOV-2001  
Pan troglodytes DNA, clone: PTB-105C09.R, genomic survey sequence.

AG101978.1 GI:16722495

GSS.

Pan troglodytes (chimpanzee)

Pan troglodytes

Pan troglodytes

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

1

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Unpublished

BAC end sequences of Library PTB

2 (bases 1 to 717)

Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,

Totoki, Y., Watanabe, H. and Sakaki, Y.

Direct Submission

Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan

(E-mail: chimpanzee@riken.go.jp, URL: http://hgp.gsc.riken.go.jp/;

Tel: 81-45-503-9111, Fax: 81-45-503-9170)

Clones are derived from the chimpanzee BAC library PTB This BAC end

was generated during the R&D process and may have higher chance of

```
clone tracking errors.
PRIMERS
  LIBRARY      Sequencing: M13Rev
    Vector      : pKS145
    R.Site 1    : SacI
    R.Site 2    : SacI
    Location/Qualifiers
      1. 717
    /organism="Pan troglodytes"
    /mol_type="genomic DNA"
    /db_xref="taxon:9598"
    /sex="male"
    /cell_type="lymphoblast"
    /clone_lib="PTB Chimpanzee Male BAC Library"
ORIGIN
  Query Match      2.7%; Score 241.6; DB 29; Length 717;
  Best Local Similarity 62.9%; Pred. No. 2e-53;
  Matches 409; Conservative 0; Mismatches 234; Indels 7; Gaps 2;
  QY 184 CCGTCTCTGCGGTTTACCGTTTATCTAATTACGGGAACGTGCTGGTGGGAAAGTACC 243
  DB 69 CCGTCTCTGCGGTTTACCGTTTATCTAATTACGGGAACGTGCTGGTGGGAAAGTACC 128
  QY 244 AGCATATCGGCGCTGTACCAAAATTTAACTGCTGATCATCGGCGGCGACCATAGACC 303
  DB 129 AGCGTTAGTCTGCTCCACACACGATGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 188
  QY 304 GCACAGACCTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 363
  DB 199 GCACAGACCTATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248
  QY 364 TTTAAGAGCCGACATCAATATATAGC-----CGTCAGAAAGTCTATCAGACGGAGCC 417
  DB 249 TTTAAGAGCCGACATCAATATATAGC-----CGTCAGAAAGTCTATCAGACGGAGCC 308
  QY 418 GTATCCATAGACCAATTCAGCAACAGAGCTATCGAAGTATTCGGCGGTTATAGTGCAC 477
  DB 309 GCGGCGCTTAGAGGAACCTCAGCGGCGGACCTGCGCAAACTACTGCGCGGTTCTTCCGAC 368
  QY 478 ATTATGAAGAGGTTTATGCGGAAACCAATGGCATGTACGGGAGCTATATCAACGGC 537
  DB 369 ATTGCGCGGAGTTTCAGCGCAACCAAGCCAGGGGGTCTACTCGGAGTCTCTGGCCCC 428
  QY 538 AATTTGAACCGCTCTCAGAAATGACGGACCGTGTATGGAATTCGAATATATATGTA 597
  DB 429 GCGTTTGAAGTCTCAGAAATGACGGACCGTGTATGGAATTCGAATATATATGTA 488
  QY 598 ATGACAGAGCGGAAACCGCTCTCTTACATATCTACCAACCGTGTGTCTTTTACTGG 657
  DB 489 GTGACGAGCTGGAACCGCTTCCGTGCAATCTACAGCGGTTCTTTTGTCTACTGG 548
  QY 658 TTCCTAAACAGTGGCTAAATACCGCTCTTTACCGCAAGGGCGGTTCCGTGATAGTA 717
  DB 549 TTTCTTCAACGCTGCTGCTACCGCACTCTACCGCGCGGCGCGCATTCCTGATCGTC 608
  QY 718 TCGCTCGGTTGCGCAACGACAGCAACCGCTTCCAGTCAACTTACCAACCGGAGCGCAA 777
  DB 609 TCGCTGAGCTCTCCAC-CAGACAGACGCTTTTCTGCTCTCTTTAGCAGAGACCCAG 667
  QY 778 AAGACGGAGATATCGTGGGAGAACATCTTCAATTCATGATAGGAA 827
  DB 668 GTAAACAAGATTGCGAGTGGGACAAACATCTCCTCTCTCTCTCTCTCTCTCTCTCTCT 717

RESULT 2
AG104414
LOCUS      Pan troglodytes DNA, clone: PTB-108F03.R, genomic survey sequence.
ACCESSION AG104414
VERSION   AG104414.1 GI:16724932
```

```

Db      482 GTGACGAAGCTGGACAGCGTTTCCTGGTGCACATCTTAACAGCGTGGTCTTTTGTACTGG 541
Qy      658 TTCTTAACAGCTGGCTGTAATATACCCCTCTTTTACCCGCCAGGGCGGTTCCGTGCATAGTA 717
Db      542 TTCTTAACAGCTGGCTGGTACCCACCTTACCGCGCGGGCGCATTTCCCTGCATCGTC 601
Qy      718 TCGTCCGTTGGCAACGAGACAAACGGTTCCAGTCAACTTACAACCGGAGCGCA 776
Db      602 TCGTGGGCTCTCCACCCAGACAGACGCTTTTCACTGCTCTCTTTAGCCACGAGACCCA 660

RESULT 3
AG126121
LOCUS      Pan troglodytes DNA, clone: PTB-136J09.F, genomic survey sequence.
AG126121
ACCESSION AG126121
VERSION   AG126121.1 GI:16655286
KEYWORDS  GSS.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
REFERENCE 1
AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
          Totoki,Y., Watanabe,H. and Sakaki,Y.
          BAC end sequences of Library PTB
          Unpublished
          2 (bases 1 to 688)
          Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
          Totoki,Y., Watanabe,H. and Sakaki,Y.
          Direct Submission
          Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22, Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
          Clones are derived from the chimpanzee BAC library PTB This BAC end
          was generated during the R&D process and may have higher chance of
          clone tracking errors.
          PRIMERS
          Sequencing: -21M13
          LIBRARY
          Vector      : pKS145
          R.Site 1    : SacI
          R.Site 2    : SacI.
          Location/Qualifiers
          1..688
             /organism="Pan troglodytes"
             /mol_type="genomic DNA"
             /db_xref="taxon:9598"
             /clone="PTB-136J09.F"
             /sex="male"
             /cell_type="lymphoblast"
             /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES             source
Query Match          2.3%; Score 213; DB 29; Length 688;
Best Local Similarity 62.1%; Pred. No. 1.2e-45;
Matches 370; Conservative 0; Mismatches 218; Indels 8; Gaps 2;

Qy      184 CCGTTCCTGCGTTTACGGTTTATCTAATTACGGGAACTGCTGTGCTGGGAAAAGTACC 243
Db      94  CCGTTCCTGCGCTTTTACGGCTATGTAAATTACGGGAAACGGGGGGTGGCAAGACACC 153
Qy      244 AGCATATCGGCGCTGTACCAAAATTTAACTGCTGTATCGGCGCGGACCAACCATAGCC 303
Db      154 AGCGTTAGTCTCCACCAACGACGACTGCTGCTGCTCGGAGGCCAACCGTGGCC 213
Qy      304 GCACAGAACCTATCGCTGGCTTAAGAGCTTCTGTCCAGATCTTCAGCGCTTTGGC 363
Db      214 GCACAGAACCTCTCCACACTCCGAGCCTACTGCCCAACCGCTATAGCGCTTTGGG 273
Qy      364 TTTAAGAGCCGACACATCAATAT-----AGCCGTCAGAAAAGCTCATCAGACCGGAGCC 417

```

```

Db      274 TTCAAGAGCCGCCACATAATATGATGTAGCGGTGACGAGTCATGTCCTCTACGGAC 333
Qy      418 GTATCCATAGAGCAAAATTCAGCAACAGGAGCTATCGAAGTATTTGGCCGGTTATAGTGGAC 477
Db      334 GCCGCCCTAGAGGAATCTCCAGCGCGGAGCTCGGCCAAATACCTGGCCGGTACTCTCCGAC 393
Qy      478 ATTATGAAGAGGTTATGGCGAAAAAACCCAAATGGCATGTACGGGACTATATCCAAACGCG 537
Db      394 ATTCCGCCCGAGTTACGGCGAAACCAAGCCAGNGGGCTCTACTCGGGAGTCTCTGGCCCC 453
Qy      538 AATTTGAAACCCCTCTCGAGAATGACCGGACCGGTGTTTATGAGCTTCCATATATTGTA 597
Db      454 GCCTTTGAGGTCTCTGAGAGACATGACCAAGNGGCGAGCTATGACCACCAATGTGATCGTG 513
Qy      598 ATCGACGAGCGCGAAACCCCTCTCTTACATATCTACCAACCGTCGCTGTTCTTTTACTGG 657
Db      514 GTGACGAAGCTGGAACGCTTTTCGTGCACATNCTAACAGCCGCTGTTCTTTGCTACTGG 573
Qy      658 TTCTTAACAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
Db      574 TTCTTAACAGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 631
Qy      718 TGCCTGCGTTGCGCAACGAGACAAACGCTTCCAGTCAACTTACAACCGGAGCCGAC 773
Db      632 TGCCTGCGCTCTCCACCCAGACAGCGCTTTTCAGTCTCTCTTTAGCCACGAGAC 687

RESULT 4
AG051826/c
LOCUS      Pan troglodytes DNA, clone: PTB-033P12.R, genomic survey sequence.
AG051826
ACCESSION AG051826
VERSION   AG051826.1 GI:16589269
KEYWORDS  GSS.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
REFERENCE 1
AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
          Totoki,Y., Watanabe,H. and Sakaki,Y.
          BAC end sequences of Library PTB
          Unpublished
          2 (bases 1 to 721)
          Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
          Totoki,Y., Watanabe,H. and Sakaki,Y.
          Direct Submission
          Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22, Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
          Clones are derived from the chimpanzee BAC library PTB This BAC end
          was generated during the R&D process and may have higher chance of
          clone tracking errors.
          PRIMERS
          Sequencing: M13Rev
          LIBRARY
          Vector      : pKS145
          R.Site 1    : SacI
          R.Site 2    : SacI.
          Location/Qualifiers
          1..721
             /organism="Pan troglodytes"
             /mol_type="genomic DNA"
             /db_xref="taxon:9598"
             /clone="PTB-033P12.R"
             /sex="male"
             /cell_type="lymphoblast"
             /clone_lib="PTB Chimpanzee Male BAC Library"

FEATURES             source
Query Match          2.3%; Score 213; DB 29; Length 688;
Best Local Similarity 62.1%; Pred. No. 1.2e-45;
Matches 370; Conservative 0; Mismatches 218; Indels 8; Gaps 2;

Qy      184 CCGTTCCTGCGTTTACGGTTTATCTAATTACGGGAACTGCTGTGCTGGGAAAAGTACC 243
Db      94  CCGTTCCTGCGCTTTTACGGCTATGTAAATTACGGGAAACGGGGGGTGGCAAGACACC 153
Qy      244 AGCATATCGGCGCTGTACCAAAATTTAACTGCTGTATCGGCGCGGACCAACCATAGCC 303
Db      154 AGCGTTAGTCTCCACCAACGACGACTGCTGCTGCTCGGAGGCCAACCGTGGCC 213
Qy      304 GCACAGAACCTATCGCTGGCTTAAGAGCTTCTGTCCAGATCTTCAGCGCTTTGGC 363
Db      214 GCACAGAACCTCTCCACACTCCGAGCCTACTGCCCAACCGCTATAGCGCTTTGGG 273
Qy      364 TTTAAGAGCCGACACATCAATAT-----AGCCGTCAGAAAAGCTCATCAGACCGGAGCC 417

```

Query Match 2.1%; Score 194.4; DB 29; Length 721;  
 Best Local Similarity 62.2%; Pred. No. 1.3e-40; Indels 0; Gaps 0;  
 Matches 306; Conservative 0; Mismatches 186; Indels 0; Gaps 0;

444 GGAGCTATCGAGTATGGCGGTATAGTGGACATTAAGAAAGAGTTATGGCGAAAAA 503  
 599 GGAGCTGGCCAAATAGTGGCGGTACTCCGACATTCGCGCGAGTTGAGCGAACAA 540  
 504 ACCCAATGGCATGTACGGGACTATATCCAAACGGGATTTTGAACCTCTCGAATGAC 563  
 539 GCTCAGAGGCTCTACTCGGAGTCTCTGGCCCGCTTTGAGTCTGAGAAATGCA 480  
 564 CGGACCTGTTTATGACATTCATATTAATGATATGATGAGAGCGGAGCCCTGTCTC 623  
 479 CCAGGGCAGCTATGGACCAACCAATGTATGCTGGTGGACGAGTGAAGCTTCCGT 420  
 624 TTACATCTACCAACCGTCTGTGTTCTTTACTGGTTCCTTAAACAGCTGGCTAAATACCC 683  
 419 GCATCTTAACAGCGGTGCTCTTTACTACTGTTCTTCAAGCGCTGATGCGTACCC 360  
 684 TCTTACCGCAGGGCGGTTCCTGCGATAGTATGCTGCTGCGCAACGAGACAA 743  
 359 ACTCTACCGCGGGCGGATTCCTGTCATGCTGCTGCGGTGCTCTCCACCCAGACGA 300  
 744 CGGTTCCAGTCAACTTACAAACACCGGAGCGCAAAAGAGCGGAGATATGCTGCGGAA 803  
 299 CGCCTTTCAGTCTCTCTTTAGCCACGAGACCCAGGTAAACAGATTTCGAGTGGACAA 240  
 804 CATCTTAACATTCATGATGGAGAGAGTCTGATCTGATGATGATGATGATGAGAGAA 863  
 239 CATCTCACTTCTGTTGGGCAACCCAGGCGCGACCTACGTTGAGCGTGGCCGCA 180  
 364 CTGGGCGCTCTTATAAACAAACGCGCTGACCGATCTGACGATTTGGACACCTTTAAA 923  
 179 CTGGGCGCTCTCATCAACAAACGCGCTGACCGAGCTGACGATTTGGACACCTCATGA 120  
 924 AATTTAGATA 935  
 119 GACCTCGAGTA 108

RESULT 5  
 AG124918  
 LOCUS  
 DEFINITION Pan troglodytes DNA, clone: PTB-135A07.R, genomic survey sequence.  
 ACCESSION AG124918  
 VERSION AG124918.1 GI:16654083  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y.  
 BAC end sequences of Library PTB  
 Unpublished  
 2 (bases 1 to 685)  
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y.  
 Direct Submission  
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22, Suenro-cho, Tsukuba, Ibaraki, 305-8565, Japan  
 (E-mail:chibbes@sc.riken.go.jp; URL:http://hgp.gsc.riken.go.jp/; Tel:81-45-503-9111; Fax:81-45-503-9170)  
 Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.  
 PRIMERS  
 Sequencing: M13Rev  
 LIBRARY Vector : PKS145

R.Site 1 : SacI  
 R.Site 2 : SacI.  
 Location/Qualifiers  
 1..685  
 /organism="Pan troglodytes"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9598"  
 /clone="PTB-135A07.R"  
 /sex="male"  
 /cell\_type="lymphoblast"  
 /clone\_lib="PTB Chimpanzee Male BAC Library"

# ORIGIN

Query Match 1.5%; Score 135; DB 29; Length 685;  
 Best Local Similarity 53.9%; Pred. No. 1.8e-24;  
 Matches 322; Conservative 0; Mismatches 245; Indels 30; Gaps 1;

955 GAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1014  
 60 GATATCTTGGGCTACGTCGACCGCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 119  
 1015 GAATACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1074  
 120 CAGTACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 179  
 1075 AACCTACACATGCTCTAAACGCTA-----GGGGC 1104  
 180 ACCCTTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 239  
 1105 GATACCGAGGACACGAGCTCTTTACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1164  
 240 GATGGGGCGGGGTGACCATGTTTACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 299  
 1165 TTGAGGAAATACAGCGGCGCTCAACCTCAACCTCAACCTCAACCTCAACCTCAACCTCA 1224  
 300 CTGGCCGATACAGACCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 359  
 1225 AAAACCTCTTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTA 1284  
 360 AAGATTTATGACGCTCGGTAACTACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 419  
 1285 GCCACGGAATCCACTGAACTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1344  
 420 GGCACCGAGGACGAGGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 479  
 1345 AGCCACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1404  
 480 AGCTCGGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 539  
 1405 TTTGAGTTCAAAAGATCTTAGACGAGCTCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1464  
 540 GGGGACTTTATGGACACGCTCGAGGCGGACGCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 599  
 1465 CCGGAATACGTATACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1521  
 600 CCGGATGATGCTACAGCTTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 656

# RESULT 6

AG051825  
 LOCUS  
 DEFINITION Pan troglodytes DNA, clone: PTB-033P12.P, genomic survey sequence.  
 ACCESSION AG051825  
 VERSION AG051825.1 GI:16589268  
 KEYWORDS GSS.  
 SOURCE Pan troglodytes (chimpanzee)  
 ORGANISM Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H., and Sakaki, Y.  
 BAC end sequences of Library PTB

JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 656)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Tokoki, Y., Watanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chimps@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/,  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI  
Location/Qualifiers  
1. .656  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-033P12.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"  
ORIGIN  
Query Match 1.2%; Score 111; DB 29; Length 656;  
Best Local Similarity 53.5%; Pred. No. 5.7e-18;  
Matches 300; Conservative 0; Mismatches 230; Indels 31; Gaps 2;  
QY 955 GAAGTCATGAGTTACGTPAGACAGATTGTCGTTCTCTAAAGTAAGATTATGATCCTTTA 1014  
Db 90 GACATCCTGGCTACGTCGACCGCTTCGTCGTCGCCAGCGGCCCAATATGACCCCGCG 149  
QY 1015 GATACATTCGTCGACCGACCTCTTTTGTACATAGCGAGTAAGGGGTATCTAACA 1074  
Db 150 CAGTACGTGGGTGACCGCGCTCTTCCTCTCCACGCCGAGTCAAGACTTCTCTCACC 209  
QY 1075 AACCTACACATGCTCTAAGCTA-----GGGGG 1104  
Db 210 ACCCTCATGCGAGCTCAAGACTCGCGGCGAGGGGGTCCCGGGGGAACCGGGGA 269  
QY 1105 GATACAGGACACGAAGCTTTTACCTGTCCTGCGTGTGTCGAGGTGTTTGAAGCCG 1164  
Db 270 GATGGGCGGGGTGACCATGTTTACCTGCGCGGTGCGAGGTCTTCTCGATCCC 329  
QY 1165 TTTGAGGATACAGCGGCGCTCACTCAACCTCACCTGACCGAGTGGGTGACA 1224  
Db 330 CTGGCCAGATACAGACCCTGGTTCGGCTTCGGGCTTCAGGCCACACCTGGCTCCA 389  
QY 1225 AAAAACCTCTTTAAGTTAAGTAATTTCGACGTTTGTGACCGAGGACATGTCATGTC 1284  
Db 390 AAGAAATTATGACGCCCTCGGTAACTACTCGCAGTTTCGCCGACCAAGACATGTC 449  
QY 1285 GCCACGGA-ATCCACTGAAGCTTCGACCCAGTTTACCTTTATCACCAGTTTGTAAAA 1343  
Db 450 GGCACCGAGGACGAGAACGCGGTGAAAGGTCAACCTACATGTCACCTACGTCAGCA 509  
QY 1344 CAGCCAGTGTCTTTAAATGAAAAAACAATAATGATATGTGGTTTCAGGGTACGTA 1403  
Db 510 CAGCTCGGTGTCGTCACCTCAAGACCAAAAGTCCATCTGTGGCTATACCGGCACCT 569  
QY 1404 CTTTGAGTTCAAAAGATTCCTAGACAGGAACTCTTGTGGAGACGATTCGCAAGCCG 1463  
Db 570 CGGGGACTTTTGGACACGCTCGAGGCGGACAGCTTCGGAGGCCCAACCGGATGAGCA 629  
QY 1464 TCCGGAATACGTATACGGTTT 1484  
Db 630 GTCCGAGTATGCTACAGCTT 650

RESULT 7  
AJ395572/C  
LOCUS AJ395572 dkf426 Gallus gallus cDNA clone 22p14r1, mRNA sequence.  
DEFINITION AJ395572 dkf426 Gallus gallus cDNA clone 22p14r1, mRNA sequence.  
ACCESSION AJ395572  
VERSION AJ395572.1 GI:7126625  
KEYWORDS EST.  
SOURCE Gallus gallus (chicken)  
ORGANISM Gallus gallus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
Phasianinae; Gallus.  
1 (bases 1 to 696)  
REFERENCE 1 (bases 1 to 696)  
AUTHORS Plachy, J., Korn, B. and Buerstedde, J. M.  
TITLE A large database of chicken bursal ESTs as a resource for the  
analysis of vertebrate gene function  
JOURNAL Genome Res. 10 (12), 2062-2069 (2000)  
MEDLINE 20568495  
PUBMED 11116100  
COMMENT Contact: Buerstedde JM  
Cellular Immunology  
Heinrich-Pette-Institute  
Martinistr. 52, 20251 Hamburg, Germany  
Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.  
Location/Qualifiers  
1. .696  
/organism="Gallus gallus"  
/mol\_type="mRNA"  
/strain="CB"  
/db\_xref="taxon:9031"  
/clone="22p14r1"  
/tissue\_type="Bursa of Fabricius"  
/cell\_type="bursal lymphocyte"  
/dev\_stage="2-3 weeks old"  
/clone\_lib="dkf426"  
ORIGIN  
Query Match 1.2%; Score 105.8; DB 9; Length 696;  
Best Local Similarity 51.4%; Pred. No. 1.5e-16;  
Matches 270; Conservative 0; Mismatches 25; Indels 3; Gaps 1;  
QY 3634 AGCCAGGCGAGACGATGGCGCGCTTCAGGACGAGATGTCTCTCCGTTTATCAGCGT 3693  
Db 601 AGAGAAAGGGGAGGGTGAACCGCTCGACGACGTGGTGGCTTCTCTGTCATAGAGCT 542  
QY 3694 CGCCTTGTCTATAGCTTCTTCCCGACAGATATAAAGCGACCTCTGAAGCTTGGCAGA 3753  
Db 541 GCGCTTCTTGTGCTAAGCTCCCGACAGATGAAACCAACCCCATCGAGTCTTGT 482  
QY 3754 AAGGCGAGCTTATATGTAGCTGGTGAACCAATCCCGACCAAGTTGGCGTGTGACCCG 3813  
Db 481 CAGCCAGGAGACGACCATCCGTGAATCTCCAGCCCTTCTCTTGTGGAGCTGGC 422  
QY 3814 CTTCCCTCTTTCACCGTCAGATGTTTAAAGAGCAACTCCCGTTTGGCCACA 3873  
Db 421 CTGTGGGCTCGACCGTGGAGACGCGCTTGAGCAGGACACCGCTCTTGGCCAGCC 362  
QY 3874 GTCTAAGCACCGCTGAGAGGCGCTGAACCCAGGTACAGTATTGGCTATCTCTTTAA 3933  
Db 361 ACTCAGATCGCGCTGACCGAGATGGTGAAGTCTTCATATACGTCGACAGTCTTTGTA 302  
QY 3934 AATATTTTGGAGCTTGGAGGTATATCTGTAGTCCGAGCAACGCTAAAGGCTAGTCCG 3993  
Db 301 AATGTTTCCAACTGGGGGAGGTGGAACAGGTTTCTGGACACTGAAACAGAGCCCAT 242  
QY 3994 GCGCTG---ACCGCATGTAAGGATCTGGCCAGGATCAGACTTTTAAATCTCCCG 4050  
Db 241 AGCTTGATTAGGTCCATAGATAGGATCTTGTCCGAGATTCAATTTTCACATCCCTGAT 182  
QY 4051 TTCCGAACAATACGACCAACATAATTTCTATCTATCGGTGGATAAATCAGGTTACGTT 4110

```

Db      181  GTCCACATCTGTGTCCAGTGAAGACCTGCTCCGGGGGGGTACACCGTGTACCTCTT 122
Qy      4111  GTCATATCCATTAACGGCGCTTACGACAGCGCGGAGCTTCTGTTT 4155
Db      121  CCTCTCCTCGCCACGAATGCCATCAGCTCCACGAAGTAGGGCTT 77

```

```

RESULT 8
LOCUS   BU444447/c
DEFINITION 603213237f1 CSOREN13 Gallus gallus cDNA clone CHEST195j9 5', mRNA
sequence.
ACCESSION BU444447
VERSION   BU444447.1 GI:25933758
KEYWORDS EST
SOURCE   Gallus gallus (chicken)
ORGANISM Gallus gallus

```

```

REFERENCE 1 (bases 1 to 911)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection of Chicken cDNAs
MEDLINE Curr. Biol. 12 (22), 1965-1969 (2002)
PUBMED 22335534
COMMENT 12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

```

```

FEATURES
source
1..911
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer"
/db_xref="taxon:9031"
/clone="CHEST195j9"
/sex="Female"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSOREN13"
/note="Organ: ovary; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research
(1996) 6: 791, except that a significantly longer
reannealing hybridization was used."

```

```

ORIGIN
Query Match 1.1%; Score 103.6; DB 13; Length 911;
Best Local Similarity 51.1%; Pred. No. 6.8e-16;
Matches 268; Conservative 0; Mismatches 254; Indels 2; Gaps 1;
Qy 3634 AGCCAGGGGAGACGATCGCGCGCTTGAGGACGAGATGTTCTGTCGGTTATCAGCGT 3693
Db 693 AGAGAAGGGGAGGAGGTTGAACCGTCTGCAGGACGTGTGGTGGCTTCTCTGTCAATAGAGCT 634
Qy 3694 CGCCTGTCTATAGCTTCTTCCACAGCATAAACGACCTGTGAGCTTGGCAGA 3753

```

```

Db      633  GCCCTTCTTGTGCTATAAGCTCCCCACAGCATGAAACACACCCCATGCAAGTTCTTGT 574
Qy      3754  AAGCAGCTTATTATGTAGTGGTGAACCAATCCAGCCAAAGTTGGCGTGTGACCCCGC 3813
Db      573  CAGCAGGAGAGACCAATCCGTGAAGTGTCTCCAGCCCTTCTCCTTGTGGAGCTGGC 514
Qy      3814  CTTCCCTCTTTCCACCGTCAAGTGGTGTAAAGCAGAACTCCCGTTTGGCCCAACA 3873
Db      513  CTGGTGGGCTCGCACCCGCTGAGGACGGGCTTGAGCAGGAGCACACCGCTGCTTGGCCAGCC 454
Qy      3874  GTCATAGCACCGCTGAGAGGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAA 3933
Db      453  AGTCAGATCGCGGTGACCCAGGATGGGTGAAGTCTTCAATATCAGTCGACACTTTTGTGA 394
Qy      3934  AATATTTTGA--GGCTTTGGAGGTATAGTGTAGTCCGAGCAACGCTAAAGCGTAGTCCG 3991
Db      393  AATGTTTTCCAACTGGGGGAGGTGGAACAGGTTTCTGACACTGAAACAGAGCCCATGA 334
Qy      3992  GTGGCTTGACCGCGATGTTAGGATCTCTGGCCAAAGGATCAGACTTTAATATCTCCCGT 4051
Db      333  GCTTGATTAGGTCCATGATTAAGGATCTTGTCCAGATTACAATTTTCAATCCCTGATG 274
Qy      4052  TCGCAACATACGACCCACACATAATTCTATCTATCGGTGGATAAATCACGGTTACGTTG 4111
Db      273  TCGCACATCTGTGTCCAAAGTGAAGACCTGCTCCGGGGGGGGGTACACCGTGTACCTCTTC 214
Qy      4112  CTCATATCCATAACGGCTTTCAGCAGAGCCGCGAGCTTCTGTTT 4155
Db      213  CTCTCTCGGCACACGATGCCATCAGTCCACGAGTAGGGCTT 170

```

```

RESULT 9
LOCUS   BX706705/c
DEFINITION BX706705 XGC-tadpole Silurana tropicalis cDNA clone TTPA010a21 5',
mRNA sequence.
ACCESSION BX706705
VERSION   BX706705.1 GI:38368912
KEYWORDS EST
SOURCE   Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis

```

```

REFERENCE 1 (bases 1 to 987)
AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
JOURNAL Unpublished (2003)
COMMENT Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
TROPICALIS_SEQUENCE ID: TTPA010a21.pikSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site 1: EcoRI; Site 2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1..987
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TTPA010a21"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from sug of poly A+ RNA from tadpole

```



embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

```

ORIGIN
Query Match      1.1%; Score 103.6; DB 13; Length 987;
Best Local Similarity 50.9%; Pred. No. 7.1e-16;
Matches 272; Conservative 0; Mismatches 259; Indels 3; Gaps 1;

QY 3631 GGTAGCCAGGGAGACGGATGCGCGCTTGAGACGAGATGCTCTGCGGTTATACAG 3690
DB 881 GGTGAGAGGGAGAGGGTGCACAGTTTGGAGAACAAAGTGCTCTTCTGTGCAATAT 822
QY 3691 GGTGCGCTTCTATAGCCCTTTCTCCACACATATAAAGCCACCTCTGAGCTTGGC 3750
DB 821 GCTGCTTTCTTCTGAGCATTAAGCTCCCAACATGAAACCAACCAACCATGATGTTT 762
QY 3751 AGAAGGCGCTTATTATGATGCTGTGTAACCAATCCAGCCAAAGGTGGCGTGTACCC 3810
DB 761 GTTAGCCAGGACACACAGAGTCTGTGAATCTTCCATCCACAGTCTTGTGAGATT 702
QY 3811 CGGCTTCCCTTTCCACCGTCAAGATGGTGTGTTTAAAGCAGAACTCCCGTTTCCCA 3870
DB 701 TGCATTATGAGCTCGGACTGTGATACAGCAATGAGAGGAGGACTCTTGTCTGGCCCA 642
QY 3871 ACAGTCTAAGCACCCTGAGAGGCGGTGAACCCAGGTACAGTATTGGCTATCTTT 3930
DB 641 GCCTGTGAGATCAACATGCGCTGGTGTGAATCTTCAATATCAGTTTCCAGTTCTTT 582
QY 3931 AAAAATATTTTGGAGCTTGGAGTATPACTGTAGTCGGAGCAACGCTAAAGCTAGTCC 3990
DB 581 GTACATATTCACTAAGCTTGGAGGCGGAGCTGGTTTCTTAACGCTAAAGCAAGTCC 522
QY 3991 GGTGGCTTGA---CCGGAGTGAAGATCTCGGCAAGATCAGCACTTATATATCCCTC 4047
DB 521 ATGAGCTTGTATTTGGTCCATGATAAGGATCTTGACCCAAATGACAACTTCAATCTT 462
QY 4048 CGGTTCGCAACAATACACCAACCAATATCTATCTATCGTGGATATAATACAGTTAC 4107
DB 461 GATGTCACCAATTGTGTCCAGGTGAATACCTTCTTGGGGTGGATAAACTGTATTT 402
QY 4108 GTTCTCATATCCATACGGCTTCAGCAGCGCGGAGCTTCTGTTTAAAAA 4161
DB 401 CTTTCTCTCTCTGCAAAATTCGATAGCTTAAACAAAGTAGGGCTTAGCAAA 348

```

```

RESULT 10
BX719507/c
LOCUS BX719507 XGC-tadpole Silurana tropicalis cDNA clone TTPA030n15 5',
DEFINITION mRNA sequence.
ACCESSION BX719507
VERSION BX719507.1 GI:38392248
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 871)
Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.
Sanger/Xenopus tropicalis EST project 2001 (11_2003)
Unpublished (2003)
Contact: Croning MDR
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger/Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE ID: TTPA030n15.plkSP6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Nigel Garrett.
cDNA was oligo dt primed from Sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107 with

```

```

EcoRI at the 5' end and NotI at the 3' end.
Vector: pCS107; Site1: EcoRI; Site2: NotI
Host: Escherichia coli DH10B.
Location/Qualifiers
1. 871
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clones="TTPA030n15"
/dev_stage="tadpole (stage 35-40)"
/lab_host="E. coli DH10B"
/clone_lib="XGC-tadpole"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from Sug of poly A+ RNA from tadpole
embryos. EcoRI-NotI cut cDNA was then ligated into pCS107
with EcoRI at the 5' end and NotI at the 3' end"

```

```

ORIGIN
Query Match      1.1%; Score 103.2; DB 13; Length 871;
Best Local Similarity 50.7%; Pred. No. 8.6e-16;
Matches 271; Conservative 0; Mismatches 260; Indels 3; Gaps 1;

QY 3631 GGTAGCCAGGGAGACGGATGCGCGCTTGAGACGAGATGCTCTGCGGTTATACAG 3690
DB 864 GGTGAGAGGGAGAGGGTGCACAGTTTNGAGAACCAAGGTGCTCTTCTGTCAATATT 805
QY 3691 GGTGCGCTTCTATAGCTTTCTTCCACAGATATAAAGCCACCTCTGAAGCTTGGC 3750
DB 804 GCTGCTTTCTTCTGAGCATTAAGCTTCCCAACATGAACCAACCAACCATGCTTTT 745
QY 3751 AGAAGGCGCTTATTATGATGCTGTGTAACCAATCCAGCCAAAGGTGGCGTGTACCC 3810
DB 744 GTTAGCCAGGACACCAAGAGTCTGTGAATCTTCCATCCACAGTCTTGTGAGATT 685
QY 3811 GCCTTCTCTCTTCCACCGTCAAGTGTGTTTAAAGCAGAACTCCCGTTTGGCCCA 3870
DB 684 TGCATTATGAGCTCGGATCTGTGATACAGCAATGAGAGGAGGACTCTTGTGTCGCCA 625
QY 3871 ACAGTCTAAGCACCCTGAGAGGCGGTGAACCCAGGTACAGTATTGGCTATCTTT 3930
DB 624 GCTGCTCAGATCACCATGCTGCTGGGTGTAATACTCTTCAATATCAGTTTCCAGTTCTTT 565
QY 3931 AAAAATATTTTGGAGCTTGGAGGATPACTGTAGTCGGAGCAACGCTAAAGGCTAGTCC 3990
DB 564 GTACATATTCACTAAGCTTGGAGGAGGGGAGCTGTTTCTTAACGCTTAAGCAAGTCC 505
QY 3991 GGTGCTTCTGA---CCGGAGTGAAGATCCTGCGCAAGGATCAGACTTTTAATATCTCTC 4047
DB 504 ATGTCCTTGTATTTGGTCCATGATAGGATCTTGACCCAAATGACAACTTCAATCTTT 445
QY 4048 CGGTTCGCAACAATACAGCACCACATATCTATCTATCGTGGATATAATCAGGTTAC 4107
DB 444 GATGTCACCAATTGTGTCAGGTTGAATACCTTCTGCGGGTGGATATAAAGCTGTATT 385
QY 4108 GTTGTCTCATATCCATAAGCGCTTCCAGCAGAGCCCGAGCTTCTGTTTAAAAA 4161
DB 384 CTTTCTCTCTCTGCAAAATTCGATAGCTTAAACAAAGTAGGGCTTAGCAAA 331

```

```

RESULT 11
AL894178/c
LOCUS AL894178 XGC-egg Silurana tropicalis cDNA clone Tegg092j11 5', mRNA
DEFINITION sequence.
ACCESSION AL894178
VERSION AL894178.2 GI:38694463
KEYWORDS EST.
SOURCE Silurana tropicalis (western clawed frog)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 645)

```

AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.  
TITLE Sanger Xenopus tropicalis EST project 2001 (11\_2003)  
JOURNAL Unpublished (2003)  
COMMENT On Sep 16, 2002 this sequence version replaced gi:22944729.  
Contact: Taylor R

Sanger Institute  
Hinxton, Cambridgeshire, CB10 1SA, UK  
Email: trop@sanger.ac.uk  
Sanger Xenopus tropicalis EST project 2001  
TROPICALIS\_SEQUENCE\_ID: TEGG092j11.plkSP6  
Sequencing primer: SP6  
This sequence is from a Xenopus Gene Collection (XGC) library  
constructed by Aaron M. Zorn.  
cDNA was oligo dT primed from Sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the  
5' end and NotI at the 3' end.  
Vector: pCS107; Site 1: EcoRI; Site 2: NotI  
Host: Escherichia coli XL1-blue.  
Location/Qualifiers

## FEATURES

source  
1..645  
/organism="Silurana tropicalis"  
/mol\_type="mRNA"  
/db\_xref="taxon:8364"  
/clone="TEGG092j11"  
/dev\_stage="egg"  
/lab\_host="Escherichia coli XL1-blue"  
/clone\_lib="XGC-egg"  
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA  
was oligo dT primed from Sug of poly A+ RNA from egg.  
EcoRI-NotI cut cDNA was then ligated into pCS107 with  
EcoRI at the 5' end and NotI at the 3' end"

## ORIGIN

Query Match 1.1%; Score 103; DB 9; Length 645;  
Best Local Similarity 51.0%; Pred. No. 8,5e-16;  
Matches 269; Conservative 0; Mismatches 255; Indels 3; Gaps 1;  
QY 3638 AGGGAGACGGATGGCGCGCTTGAGACGAGATGCTCTGCGTTATATACGCGTCGCC 3697  
Db 640 AGGGAGAGGGTGCACAGTTTGAGAACAGGTGCGCTCTTCTGTCATATTTGCGCT 581  
QY 3698 TTGCTATAGCTTCTCCACAGATATAAAGCGACTCTGAAGCTGGCAGAAAGG 3757  
Db 580 TTCTCTGAGATGAAGTCCCAACATATGAACCAACATGATGTTTGTGTTAGC 521  
QY 3758 CAGCTTATTATGATGGTGAACCAATCCACGCAAGGTGGCGGTGACCCGCGCTTC 3917  
Db 520 CAGGACACCAAGAGTCTGTGAAGTCTCCATCCACAGTCTCTGTGAGAGTTTGCATTA 461  
QY 3818 CCTCTTCCACGTCAGATGGTGTAAAGACGAGACTCCCGTTTGGCCACAGTCT 3877  
Db 460 TGAGCTGGAGTGTGATGATGATGAGGAGGAGTCTCTGCTTGGCCGCGCTGTC 401  
QY 3878 AAGCACCCGTGAGAGGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAATA 3937  
Db 400 AGATCACCATGCCCTGGGTGTGAATCTCTTCATATCAGTTTCCAGTCTCTTTGTACATA 341  
QY 3938 TTTTGTAGGCTTGAGGTATCTGTAGTCCGAGCAACGCTTAAGGTAGTCCGTTGGCT 3997  
Db 340 TTCCTAAGCTTGAGGAGGGGGACTGGTTTCTTAACGCTTAAGCAAGATGCTATGCTC 281  
QY 3998 TGA---CGCGATGTTAGGATCTTGCGCAAGGATCAGCACTTTAATATCTCCGTTTCG 4054  
Db 280 TGATTTGTGTCATGATAGGATCTTGACCCAAATGACACCTTCATCTTTGATGTCA 221  
QY 4055 CAACAATACGACCAACATATTTCTATCTATCGGTGATTAATACAGGTATGCTGCTC 4114  
Db 220 ACCATTTGTGCCAGGTGAATACCTCTTCTGGGGGTGATAAACGTGTATTTCTTCTC 161  
QY 4115 ATATCCATAAGCGCTTCAGACAGCGCGGAGCTTCTGTTTAAAAA 4161  
Db 160 TCTTCTGCACAAATTCATAGCTTAACAAGTAGGCTTAGCAAA 114

## RESULT 12

CF521628/c

LOCUS

DEFINITION

CF521628 699 bp mRNA linear EST 10-SEP-2003  
AGENCOURT 15508682 NICHD\_XGC\_Kid1 Xenopus laevis cDNA clone  
IMAGE:7012130 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis (African clawed frog)

Xenopus laevis

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Igor Dawid

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLM14721 row: m column: 24

High quality sequence stop: 634.

Location/Qualifiers

1..699

/organism="Xenopus laevis"

/mol\_type="mRNA"

/db\_xref="taxon:8355"

/clone="IMAGE:7012130"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NICHD\_XGC\_Kid1"

/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 2.2 kb. Constructed by Life

Technologies. Note: This is a Xenopus Gene Collection

(XGC) library."

ORIGIN

Query Match

Best Local Similarity

Matches

272; Conservative

0; Mismatches

261; Indels

3; Gaps

1;  
QY 3638 AGGGAGACGGATGGCGCGCTTGAGACGAGATGCTCTGCGTTATATACGCGTCGCC 3697  
Db 613 AGAGAGAGGGTGCACGCTTTGAAGAACATGGTGGCTCTTCTGCTCAATATTGCTGCT 554  
QY 3698 TTGCTATAGCTTTCTTCCCCACAGCATATAAAGCGACCTCTGAAGCTTGGCAGAAAG 3757  
Db 553 TTCTTCTGGCATAGCTCTCCACAGCATGAAACCAACCCATCCATGTTCTTATTAGC 494  
QY 3758 CAGCTTATTATAGCTGTGAACCAATCCACGCAAGGTGGCGGTGACCCCGCTTC 3817  
Db 493 CAGGATACCAACAACATCTGTGAACCTGTTCCCATCCACAGTCTTGTGAGAGTTGCATTA 434  
QY 3818 CCTCTTTCCACCGTCAGAAATGGTGTAAAGACAGAACTCCCGTTTTGCCAACAGTCT 3877  
Db 433 TGAGCTCGACCTGTGACACAGCATTTAGGAGGAGGACTCTCTTTGTGGCCAGCTGTG 374  
QY 3878 AAGACCCCGTGAAGAGCGGTGAACCCAGGTACAGTATTGGCTATCTCTTTAAATA 3937  
Db 373 AGATCCCATGATGACCTGGCGGTGAAATCTTCGATATCGGTTTCTAGTTCTTTGTACATA 314  
QY 3938 TTTTGGAGGCTTGGAGGTATAGTGTAGTCCGAGCAACGCTAAAGGCTAGTCCGCTGGCT 3997

|      |  |      |
|------|--|------|
| 313  | TTCACTAAGCTGGGTGGAGGTGGGACTGCTCTTTGCACCTAAAGCAAGACCATGAGCT     | 254  |
| 3998 | TGA---CCGGAGTGTGAAGATCTCTGGCCAAAGATCAAGACCTTTAAATATCTCTCGGTTGC | 4054 |
| 253  | TGATTGGTCCATGATAAGGATCTTGGCCATAAAATAACAACCTTTACATCTTTGATGTCA   | 194  |
| 4055 | CAACAATACGACACCAACATAATCTATCTATCGGTGGATAAATCAACGGTTACGTTGCTC   | 4114 |
| 193  | ACCATTGTGTCCAGGTGAATCACTCTTCTGGTGGTGAATAAAGTGTCATTTCTTTCTC     | 134  |
| 4115 | ATATCCATACGGCGCTTCCACAGACCGCGAGCTCTTGTTTTAAATGGAGAAAG          | 4170 |
| 133  | TCCTCTCGAATGAATAATTGATCTGTGACAAAGTAGGGTTTGGCAACTCAGCAAG        | 78   |

|            |            |       |          |                                 |      |           |                                |
|------------|------------|-------|----------|---------------------------------|------|-----------|--------------------------------|
| RESULT 13  | AL669156/c | LOCUS | AL669156 | 806 bp                          | mRNA | linear    | EST 14-JAN-2002                |
| DEFINITION |            |       | AL669156 | directional larval cDNA library |      |           | <i>Ciona intestinalis</i> cDNA |
|            |            |       |          | clone 043ZC09 5',               | mRNA | sequence. |                                |

AL569156 GI:18142413  
 EST.  
 AL569156.1  
 Ciona intestinalis  
 Ciona intestinalis  
 Ciona intestinalis  
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;  
 Phlebobranchia; Cionidae; Ciona.  
 1 (bases 1 to 806)  
 Genoscope.  
 Ciona intestinalis directional larval cDNA library  
 Unpublished (2002)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [segre@genoscope.cns.fr](mailto:segre@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 IMPORTANT: this sequence may contain errors. The cDNA intestinalis  
 library from which the clone was isolated may be contaminated with  
 cDNAs from bacteria or other Eukarya.  
 Directional larval cDNA library originate from Dr.M.Branno,  
 Stazione A.Dohrn, Naples, Italy, and was prepared in  
 pBluescript2SK+.

```
FEATURES
Location/Qualifiers
1..806
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/cdate="0432C09"
/clone.lib="directional larval cDNA library"
/note="vector: pBluescript2Sx+"

```

| ORIGIN | Query Match           | 1.1%;   | Score 101;         | DB 9;     | Length 806; |
|--------|-----------------------|---|--------------------|-----------|-------------|
|        | Best Local Similarity | 54.5%;  | Pred. No. 3.3e-15; |           |             |
|        | Matches 242;          | Conservative 0;   | Mismatches 198;    | Indels 4; | Gaps 2;     |
| Qy     | 3658                  | CTTGAGGACGAGATGCTCTCTCCGTTATCAGGGTCGGCTTGCTATAGCCCTTCTTCC     | 3717               |           |             |
| Db     | 729                   | CTTCAGCACAGGTGACCTTGCTGGTCTATTTTGTCTCCCTTCCTGTGACGATAACTNCC   | 670                |           |             |
| Qy     | 3718                  | CCACAGCATAAAAAC-GCACCTCTGAAGCTTGGCAGAAAGCGAGCTTATTATGTAGCTGG  | 3776               |           |             |
| Db     | 669                   | CCACAGCATAAACACAAACCCCGTTAAGTTGGTGTGATCCATGAAATGACGGAATCTG    | 610                |           |             |
| Qy     | 3777                  | TGAACCAATGCCAGCCAAAGGTGGCGGTGACCCCGCTTCCCTCTTTCACCGGTCAAGAA   | 3836               |           |             |
| Db     | 609                   | TCAACTTCTCCCATCCCTTATCTTATGTGAGTTAGGTTCCACCATCTCTTACTGTTAACA  | 550                |           |             |
| Qy     | 3837                  | TGCTGTTTTAAAAGCAGAACTCCCGTTTTCGCCCAACAGTCTAAGCACCCGTCAGAAAGAG | 3896               |           |             |
| Db     | 549                   | CGCATTGAGACGAGAACACCCCTGATCAGCCCAACCCCTGTTAAATCTCCATGCTGGG    | 490                |           |             |
| Qy     | 3897                  | CGGTGAACCCAGGTCAGTATGGCTATCTCTTTAAAAAATATTTTTGAGGCTTGGAGTA    | 3950               |           |             |

|      |  |      |
|------|--|------|
| 489  | TATTGAATCCTTCAATATACATTCATCTTCAGTTCITTTGAACATGTTNTTAACTCGGTGGAG  | 430  |
|      |  |      |
| 3957 | TACTGTAGTCCGGAGCAACGCTAAAGGCTAGTCCGGTGGCTTGACC---CGCATGTTAAG     | 4013 |
|      |  |      |
| 429  | TACGTATGCTTCTTCTGAACACTAAAGCATAAACCATGAGCTTGTCCAGGCCCATGGTAAAG   | 370  |
|      |  |      |
| 4014 | GATCTGTGGCCAAAGGATCAGACTTTTAATATCTCCCGGTTGCGCAACAATACGACCAACCACA | 4073 |
|      |  |      |
| 369  | GGTCTTGCCCAAGTATTACACTTGTATGTCCTCAATTTTACAATGTGTGTGTCATGCAA      | 310  |
|      |  |      |
| 4074 | TAATTCATCTATCGGTGGATAAA  | 4097 |
|      |  |      |
| 309  | ACACCTGTGTGTTCTGGAGGGTAAA  | 286  |
|      |  |      |

|            |            |   |        |      |        |                 |
|------------|------------|---|--------|------|--------|-----------------|
| RESULT 14  | BI940552/c | BI940552  | 574 bp | mRNA | linear | EST 18-OCT-2001 |
| LOCUS      |            |   |        |      |        |                 |
| DEFINITION |            | db71f06.y1 Wellcome CRC psk egg <i>Xenopus laevis</i> cDNA clone IMAGE:3378467 5' similar to SW:UNG HUMAN P13051 URACIL-DNA GLYCOSYLASE PRECURSOR <i>i</i> , mRNA sequence. |        |      |        |                 |

|           |  |
|-----------|--|
| ACCESSION | BI940552   |
| VERSION   | BI940552.1   |
| KEYWORDS  | GI:16255024  |
| SOURCE    | EST.   |
| ORGANISM  | Xenopus laevis (African clawed frog)   |
|           | Xenopus laevis   |
|           | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Xenopodinae; Xenopus. |

|           |   |
|-----------|---|
| REFERENCE | 1 (bases 1 to 574)  |
| AUTHORS   | Clifton, S. J. Johnson, S. L., Blumberg, B., Song, J., Hillier, L.,<br>Zaper, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y.,<br>Person, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,<br>Waterston, R., and Wilson, R. |
| TITLE     | WashU Xenopus EST project, 1999   |
| JOURNAL   | Unpublished (1999)  |
| COMMENT   | Other ESTs: db71f06.x1<br>Contact: Sandy Clifton, Ph.D.<br>WashU Xenopus EST project, 1999<br>Washington University School of Medicine<br>4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA<br>Tel: 314 286 1800<br>Fax: 314 286 1810                |

Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
Library constructed by N. Garrett, P. LeMaire, A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute). DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/[LINDL at: info@image.lindl.gov](http://LINDL.at:info@image.lindl.gov)  
Seq primer: -40RP from Gibco  
High quality sequence stop: 454.

| FEATURES   | source |
|--|--------|
| Location/Qualifiers  |        |
| 1..574   |        |
| /organism="Xenopus laevis"   |        |
| /mol_type="mRNA"   |        |
| /db_xref="taxon:8355"  |        |
| /clone="IMAGE:3378467"   |        |
| /tissue_type="egg"   |        |
| /lab_host="DH10B (phage-resistant)"  |        |
| /clone_lib="Wellcome CRC psk egg"  |        |
| /note=vector: plusecript SK-; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Library was constructed by N. Garrett, P. LeMaire A.M. Zorn, and J.B. Gurdon (Wellcome/CRC Institute). Note: This is a Xenopus Gene Collection (XGC) library." |        |

|                           | ORIGIN   |
|---------------------------|--|
| Query Match               | 1.1%; Score 100.8; DB 12; Length 574;                                |
| Best Local Similarity     | 50.6%; Pred. No. 3.2e-15;  |
| Matches 271; Conservative | 0; Mismatches 262; Indels 3; Gaps 1.                                 |
| Oy                        | 3638 AGGGAGACGCGATGGCGGCCCTTGAAGCAGAGATCTCTGTCCGTTTTATCAGCGTCGCC 369 |

Db 558 AGAGGAGGGTGCACGGTTTGAAGACATGGTGGCTTTTCTGTCATATATGCTGCT 499  
Qy 3698 TTGCTATAGCTTTCTTCCCGCAGCATATAAAGCAGCCTCTGAAGCTTGGCAGAAAG 3757  
Db 498 TTCTTCTGGCAGTAGGCTCCCGCAGCATGAAAGCAACCAACCTATGTTCTTATTTAG 439  
Qy 3758 CAGCTTATATAGCTGGTGAACCAATCCAGCAGGTTGGCGTGTACCCCGCTTC 3817  
Db 438 CAGATACCAACATCTGTGAAGTGTTCCTCCATCCACAGTCTTGTGAGAGTTGCAATTA 379  
Qy 3818 CTTCTTCCACCGTCAGAAATGGTGTATAAAGCAGAACTCCCGTTTGGCCCAACAGTCT 3877  
Db 378 TGAGCTCGCACTGTCCAGCAGCATTTGAGGAGGAGTCTCTTGTGTGGCCAGCCTGTG 319  
Qy 3878 AAGCACCGGTGAGAGGAGCGGTGAACCCAGGTGACAGTATGCTATCTCTTAAATAA 3937  
Db 318 AGATCCCATGACCTGGCGGTGAATCTTCCATATCGGTTTCTAGTTCTTTGTACATA 259  
Qy 3938 TTTTGTAGGCTTGGAGTATATCTGTAGTCGGAGCAACGCTAAAGCTAGTCGGGTGCT 3997  
Db 258 TTCACTAAGCTGGGTGGAGTGGGACTGCTTCTTGACACTAAAGCAAGACCATGAGCT 199  
Qy 3998 TGA---CCGCGATGTAAGTATCTGGCCAGGATCAGACTTATATCTCCCGTTCG 4054  
Db 198 TGATTTGGTCATGATAGGATCTTGGCCCTAAATACAACTTTTCACTTTTGATGTC 139  
Qy 4055 CAACATACGACACCAACATATCTATCTATCGGTGATAAATCAGGTTAGTGTGCTC 4114  
Db 138 ACCATTGTTGCCAGGTGAATACCTCTTGTGGTGAATAAAGTGTGCAATTTCTTCTC 79  
Qy 4115 ATATCCATAGCGCTTACAGAGCGCGAGCTTCTGTTTAAATGAGAGAG 4170  
Db 78 TCTTCTGCAATGAAATTTGATAGCTTGACAAAGTAGGGTTTGCAAACTCAGCCAG 23

RESULT 15  
CD321504/c  
LOCUS  
DEFINITION  
StrPUS38.002866 Sea urchin embryo 7hr cleavage stage cDNA library  
MPMP538 Strongylocentrotus purpuratus cDNA clone  
CALTP538D1025;MPI\_538\_25D10 5', mRNA sequence.  
CD321504  
CD321504.1 GI:34793565  
EST.  
SOURCE  
Strongylocentrotus purpuratus  
Strongylocentrotus purpuratus  
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
Echinoidea; Euechinoidea; Echinacea; Echinoida;  
Strongylocentrotidae; Strongylocentrotus.  
1 (bases 1 to 687)  
Poustka,A.J., Groth,D., Hennig,S., Thamm,S., Cameron,A., Beck,A.,  
Reinhardt,R., Herwig,R., Panopoulou,G. and Leirach,H.  
Generation, annotation, evolutionary analysis, and database  
integration of 20,000 unique sea urchin EST clusters  
Genome Res. 13 (12), 2736-2746 (2003)  
Contact: Poustka AJ  
Laboraty 145, dept Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Innestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: poustka@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting  
(ONF) to reduce sequencing redundancy. According to the ONF  
procedure, clones that display the same hybridisation matrix with a  
battery of 200 smer oligonucleotides are grouped into clusters. One  
clone per ONF cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well as the coordinates of  
the other clones assigned to the same ONF cluster as the clone from  
which the above EST is generated is available at the sea urchin  
project web site at: [http://www.molgen.mpg.de/ag\\_seaurchin/](http://www.molgen.mpg.de/ag_seaurchin/) . cDNA

clones and filters are distributed via the Resource Center/Primary  
Database of the German Human Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCAGAGGTTTACACTTATGCTTCCGGCTG 3' (M13RSP) 5'-seq  
BACKWARD: 5' GCTATTACGACGCTGGGAGAGGGATGTG 3' (M13FSP) 3'-seq  
Seq primer: 5'-CCGTCGCGAATTCGCGGT-3' pSPORT3/86  
High quality sequence stop: 687.  
Location/Qualifiers  
1. 687  
/organism="Strongylocentrotus purpuratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:7668"  
/clone="CALTP538D1025;MPI\_538\_25D10"  
/tissue type="whole embryo"  
/dev stage="embryonic 7hr"  
/lab host="E. coli, X11 blue"  
/clone\_lib="Sea urchin embryo 7hr cleavage stage cDNA  
library MPMP538"  
/note="vector: pSPORT1; Site 1: NotI; Site 2: SalI; Random  
primed and directionally cloned in pSPORT1 vector using a  
NotI (5'-PCATAGTTCTAGATCGGAGCGGCCCC (T)15-3' and a  
SalI 5'-TCGACCCACGGGTCG-3' adapters (Gibco BRL)"

FEATURES

source  
Query Match 1.1%; Score 100.8; DB 14; Length 687;  
Best Local Similarity 52.7%; Pred. No. 3.4e-15; Indels 3; Gaps 1;  
Matches 235; Conservative 0; Mismatches 208;  
Qy 3656 GCGTTGAGGAGCAGATGCTCTGTCGGTTTATCAGCGTCGCTTGTCTATAGCCTTTCTT 3715  
Db 687 GCGTTCAAGACGTGATGGCGCTGNNAGTTAATGATGCTTCTTCTCTGTGTCATAGGAT 628  
Qy 3716 CCCACACGCTATAAAGCAGCCTCTGAAGCTTGGCAGAAAGGAGCTTATATGATAGCTG 3775  
Db 627 CNCCACAGCATGANAACGACTCCATTGAGTGTGAGTTCAGCCAGGAGATGACCGCTGTG 568  
Qy 3776 GTGAACCAATCCAGCAGGTTGGCGGTGACCGCGCTTCCCTCTTTCACCGCTCAGA 3835  
Db 567 GTGAGTTTCTCNCAGCCTTGTCTTGTGTGAGTTGGTTTGGCCCTTACAGTCAAG 508  
Qy 3836 ATGGTGTTTAAAGCAGAACTCCCGCTTTTCCCAACAGTCTAAGCAGCCCGTGAAGGA 3895  
Db 507 ACTGCATTCAAAAGAACGACTCCTTGTGTTAGACCATCAATAGGTGTCCTGCTGGA 448  
Qy 3896 GCGGTGAACCCAGGTACAGTATGCTATCTCTTTAAATAATTTTGGAGCTTGGAGGT 3955  
Db 447 TGCTCAAGCCTTAAATGCTCTTGTGAGGACCTTGTACATGTTTGCAGACTTGGGGGT 388  
Qy 3956 ATACTGTAGTCCGAGAGCAACGCTAAAGCTAGTCCGGTGGCTTGACC---GCGATGGTAA 4012  
Db 387 GCGTTCACTCTCTTTTGTACACTGAAACCGTGTGCTTGACCAAGTCCATGGTAT 328  
Qy 4013 GATCTCTGGCAAGGATCAGCACTTAAATATCTTCCGTTTCGCAACATACGACCAC 4072  
Db 327 GGATCTTGACCAAGGATCACCACCTTACATCTTTGATATCAAGCCCTGTGTCATGAG 268  
Qy 4073 ATAATCTATCTATCGGTGGATAAT 4098  
Db 267 AATACCTGATGTGAGGGTGGTAGAT 242

ORIGIN

Query Match 1.1%; Score 100.8; DB 14; Length 687;  
Best Local Similarity 52.7%; Pred. No. 3.4e-15; Indels 3; Gaps 1;  
Matches 235; Conservative 0; Mismatches 208;  
Qy 3656 GCGTTGAGGAGCAGATGCTCTGTCGGTTTATCAGCGTCGCTTGTCTATAGCCTTTCTT 3715  
Db 687 GCGTTCAAGACGTGATGGCGCTGNNAGTTAATGATGCTTCTTCTCTGTGTCATAGGAT 628  
Qy 3716 CCCACACGCTATAAAGCAGCCTCTGAAGCTTGGCAGAAAGGAGCTTATATGATAGCTG 3775  
Db 627 CNCCACAGCATGANAACGACTCCATTGAGTGTGAGTTCAGCCAGGAGATGACCGCTGTG 568  
Qy 3776 GTGAACCAATCCAGCAGGTTGGCGGTGACCGCGCTTCCCTCTTTCACCGCTCAGA 3835  
Db 567 GTGAGTTTCTCNCAGCCTTGTCTTGTGTGAGTTGGTTTGGCCCTTACAGTCAAG 508  
Qy 3836 ATGGTGTTTAAAGCAGAACTCCCGCTTTTCCCAACAGTCTAAGCAGCCCGTGAAGGA 3895  
Db 507 ACTGCATTCAAAAGAACGACTCCTTGTGTTAGACCATCAATAGGTGTCCTGCTGGA 448  
Qy 3896 GCGGTGAACCCAGGTACAGTATGCTATCTCTTTAAATAATTTTGGAGCTTGGAGGT 3955  
Db 447 TGCTCAAGCCTTAAATGCTCTTGTGAGGACCTTGTACATGTTTGCAGACTTGGGGGT 388  
Qy 3956 ATACTGTAGTCCGAGAGCAACGCTAAAGCTAGTCCGGTGGCTTGACC---GCGATGGTAA 4012  
Db 387 GCGTTCACTCTCTTTTGTACACTGAAACCGTGTGCTTGACCAAGTCCATGGTAT 328  
Qy 4013 GATCTCTGGCAAGGATCAGCACTTAAATATCTTCCGTTTCGCAACATACGACCAC 4072  
Db 327 GGATCTTGACCAAGGATCACCACCTTACATCTTTGATATCAAGCCCTGTGTCATGAG 268  
Qy 4073 ATAATCTATCTATCGGTGGATAAT 4098  
Db 267 AATACCTGATGTGAGGGTGGTAGAT 242

Search completed: April 11, 2004, 07:57:51  
Job time : 13564 secs